UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN
INFORMATICA SISTEMI E TELECOMUNICAZIONI
INDIRIZZO IN TELEMATICA E SOCIETA' DELL'INFORMAZIONE

CICLO XXVIII

COORDINATORE Prof. Luigi Chisci

Twitter and codified hashtags for weather warning in Italy.

Settore Scientifico Disciplinare SPS.08

Dottorando
Dott. Valentina Grasso

Tutori
Prof. Paolo Nesi
Dott. Alfonso Crisci

Coordinatore
Prof. Luigi Chisci

Anni 2012/2017

Università degli Studi di Firenze, Dipartimento di Ingegneria
dell'Informazione (DINFO).

A mia mamma

Acknowledgments

I would like to acknowledge the efforts and inputs of my supervisor, Prof. Paolo Nesi, and all my colleagues of the LaMMA Consortium and CNR Ibimet who have been a great help during my research. In particular, my thanks go to Alfonso Crisci, my colleague at CNR Ibimet, who tutored me during the research. Without his material and immaterial support this work would have never seen the light of day.

I also would like to acknowledge the work of the many collaborators of DISIT LAB of University of Florence, in particular Simone Menabeni, Alice Cavaliere, Gianni Pantaleo, Imad Zaza who developed many of the functionalities of the TwitterVigilance platform, a fundamental tool to accomplish my research work.

I am also very grateful to both the external reviewers who read my work, Christopher Brewster and Francesca Comunello. Their suggestions and critical views gave me the opportunity to greatly improve this work.

I would like to thank my dear friend Daniela for proofreading the English manuscript. If typos are still present, the responsibility is only mine.

Finally I would like to thank my family, especially my husband and kids, because without their love and affection nothing really matters.

Contents

Chapter 1

Introduction

In this chapter, we give an overview of the context and motivations of this thesis summarizing its main contributions.

This research places itself among the several studies that investigate the use of social media in emergency situations. Social media are Internet-based applications that promote high social interaction and user generated content, both at a one-to-many or a many-to-many scale. Social media platforms like Facebook, Twitter, and Instagram, have enormously increased the amount of information exchanged every day on the Internet, especially during particular events. Emergencies are no exception, allowing the general public to play an active role during a disaster contributing to information exchange. In the last years, many research works have demonstrated the critical role played by Twitter during emergencies. When hundreds or thousands of messages are conveyed in a short period, people may feel overwhelmed by this deluge of information. In Twitter hashtags (words prefixed by the symbol #) are used as filters, to label messages and coordinate conversations, helping users to follow relevant information about specific topics. In recent years, governments and international organizations, like United Nations Office for the Coordination of Humanitarian Affairs (UN OCHA), proposed strategies to codify the use of hashtags during disaster events. A clear convention to create hashtags may improve the monitoring of information before, during and after a disaster. In Italy, a set of codified hashtags has been proposed for Twitter conversation in case of weather warnings. This research investigates the effective uptake of those hashtags in Italy. Following the proposal of a list of regional hashtags to include in tweets during weather emergencies, we have investigated if this recommendation has become a practice in Italy. To make an assessment, we realized a one-year monitoring of the proposed codified hashtags, and we analyzed collected data. The analysis showed that in some regions this practice created a hashtag-community composed of local and remote users. Features of these hashtag-communities have been analyzed and interpreted under the framework of disaster sociology theories of emergent behaviors during disasters. Emergent groups using social media and relying on codified communication practices may have a greater chance to act in coordination with institutional actors in charge of the emergency response.

Chapters are organized as follows. Chapter 2 presents a review of the literature about disaster sociology, social media and disasters, discussing major works and approaches. Chapter 3 gives an overview of the Italian weather warning system and of Twitter adoption in public organizations, also introducing the codified hashtags proposal. Chapter 4 is deputed to methodological issues, presenting the approach followed in this research. Chapters from 5 to 8 discuss the data collected for this work. Chapter 5 presents a first analysis of the codified hashtags in November 2014. Chapter 6 discusses the significant findings of one year monitoring of the whole set of codified hashtags proposed in Italy for weather warning. Chapter 7 presents some selected case studies related to high impact events. Lastly, Chapter 8 discusses the uptake of the codified hashtag during a very localized event.

This chapter is organized as follows. Section 1.1 presents a background of the research context. Section 1.2 discusses the main objective of this research and presents the elements of originality. Main contributions are discussed in section 1.3. The further section, 1.4 presents the limits of this research, elements that were not fully discussed and that need to be taken into account for further works and discussions. Last section, Definitions, presents some key terms and concepts used in this work.

1.1 Background

1.1.1 Social Media

Social media (SM) refers to a variety of web-based platforms and services which allow users to develop public or semi-public profiles and content and to connect with other users' profiles [14, 20]. Examples of social media are blogs and micro-blogs (such as Blogger, Twitter, WordPress), discussion forums (such as Reddit), digital content sharing platforms (such as Flickr, Instagram, Pinterest, YouTube), social gaming sites (such as Zynga), and social networking sites (such as Facebook, Google+, LinkedIn) [9].

Boyd and Ellison (2008) [21] identify three critical traits of social media. They include (1) constructing a public or semi-public profile within a bounded system, (2) articulating a list of other users with whom participants share a connection, and (3) viewing and traversing their list of connections and those made by others within the system. Authors also point out that the nature and nomenclature of these characteristics may be different from one platform to another. One of the first social networking sites was launched in 1997. It was called SixDegrees.com, after the idea that no person is separated by more than six degrees from another. On SixDegrees users could create profiles, invite friends and organize groups. After the invention of blogging, SNS starts to explode. In 2002 it was the time of Friendster that in one-year time gained more than three million registered users. During 2003 was launched Linkedin, which instead of creating networks of classmates and teenagers, turned to the world of professionals, becoming in few years a successful social networking platform. LinkedIn counts on more than 460 million members in 2016 (as reported by Statista.com). 2003 was also the year when MySpace was launched. It became very soon the favorite networking site for teenagers, mainly thanks to many bands and musicians entering the community. Mark Zuckerberg and his friends opened Facebook in 2004. At that time the website was conceived to connect only Harvard students. Later, the social networking platform enlarged its audience to high schools students, opening to the general public in 2006. Facebook has been ever since the most successful SNS, with more than 1.3 billion active users all over the world. In those years, Youtube and Flickr were also launched as platforms of video and picture sharing. In 2006 it was the time of Twitter, the microblogging platform, followed by Google+, which never reached the success of Facebook. With the first smartphone in 2008, social media applications and SNS exploded both in term of users and applications: Pinterest, Instagram, Snapchat, Periscope. In the crowded world of social media, Facebook is still today the number one. Used by people of all ages and nationalities to stay in touch with friends and share comments, pictures, and video. Plenty of firms, governments, public institutions and media outlet turned their attention to SNS, opening official profiles and pages to connect with users and developing entirely new public relations policies.

People use social media to connect with new people but also to maintain existing social ties [83]. The success of social media is proved by the million users worldwide. Many disciplines and scholars have investigated the reasons behind this success. Following the Users and Gratification Theory (UGT), a well-rooted media theory, users seek out media that satisfy various needs and lead to obtain certain gratification [105, 114]. These could be different depending on the media. The UGT approach has been applied also to social media. Some of the uses and gratifications related to social that scholars identified are social interaction, information seeking, pass time, entertainment, relaxation, expression of opinions, information sharing, and surveillance and watching [163, 200]. Furthermore, social networking sites may help to reinforce existing ties among people or communities by keeping their users constantly updated about what is going on. This increase in information exchange especially, among group participants, supports the reinforcement of trust among the group members potentially increasing social capital. Strengthening reciprocity and trust, social media create opportunities for civic engagement. [73].

The great variety of social media, combined with the growth in the number of users, makes evident that social media are not just a fancy but also a new and remarkable presence. These new technologies may fundamentally affect how individuals learn, interact, and organize [164]. Pew Research Center reports that 62% of American adults get news on Social Media [78]. That's why governments, organizations, corporations, NGOs are increasingly exploiting the interaction power of social media to build relationship with their stakeholders. Looking at SM users data published by Pew Research Center in November 2016, figure 1.1, it emerges that 79% of American adults use Facebook, 32% use Instagram while only 24% uses Twitter. In Europe, the Netherlands, Italy, Norway, and the UK have the highest rates of social network usage in Western Europe, with at least 69% of Internet users in each of these countries regularly visiting social networks. By contrast, France and Germany rank toward the bottom for social media usage in the region, with just 55% and 57% of Internet users, respectively, using the platforms on a monthly basis, well below the regional average of 63% (data from www.emarketer.com). This is even truer in time of disasters. In fact, while many traditional media, like newspapers and television, remain important "one-to-many" communication channel during emergencies, social media cre-

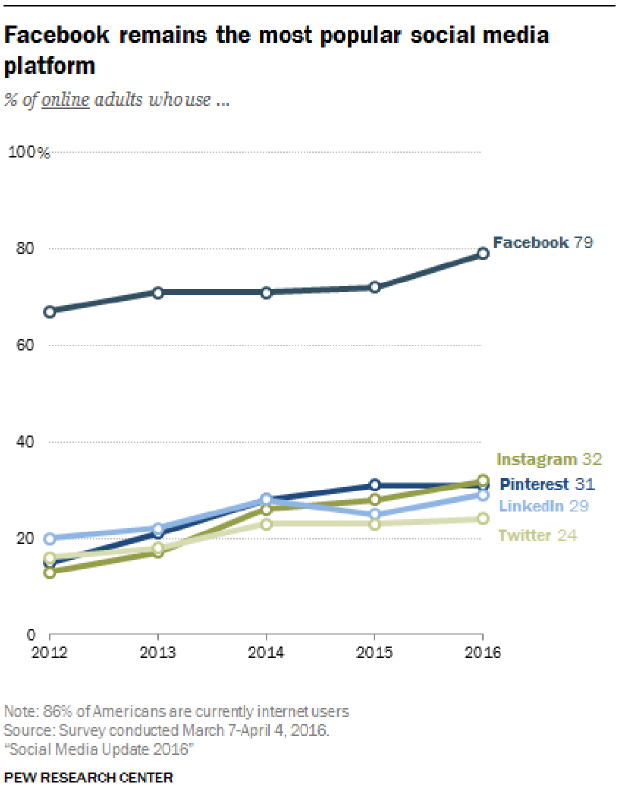


Figure 1.1: Social Media Users, by Pew Research Center, November 2016

ate opportunities for two-way interactions among the general public, media and organizations [18].

The sharing attitude pushed by social networking sites becomes even more striking during disasters. During emergencies, SM guarantees a speed of interaction and communication that is considered crucial. Studies confirm that the use of social media is considerable during crisis situations when people are in need of immediate information [185]. In particular, many studies report the use of Twitter during emergencies. Since after its launch, Twitter turned out to play a critical role during disasters and crisis. It was used to quickly share initial information and updates during the 2007 and 2008 California wildfires [93,183]. During the California wildfires, residents took pictures of the fire and reported their location on Twitter, reporting the fire's movement before journalists could get to the scene [183]. Social media functioned as important "backchannels" of communication where citizens' messages integrated the information received from institutional sources. Twitter then attracted press attention for breaking the news coverage of the catastrophic earthquake that rocked the Sichuan province of China in May 2008 [3]. In January 2010 during Haiti earthquake for the first time, Twitter showed its usefulness in a massive disaster. The earthquake was a trending topic on Twitter for weeks after the quake [167,207]. Soon after the earthquake, volunteers from all over the world used Twitter to offer translation services and other forms of help. Online volunteers used tweets to create an online map of urgent needs arising from the field [125]. It was again used during 2010 Pakistan flood, 2011 earthquake in Japan and 2011 flood in Thailand [37, 104, 110]. Still in 2011, in Australia Twitter was one of the first sources of information during the floods that hit the South East Queensland, and during the destructive earthquake in Christchurch, New Zealand, like it is reported in the work of Bruns and colleagues [23]. An outstanding example of Twitter use during a disaster has been the Hurricane Sandy of 2013, in the United States, when millions of messages went online in few days. This case is further presented in section 2.2.4.

In all these crises, social media like Twitter served as an easy-to-use technology for people seeking to share information about their status or the status of their community [191]. The news of a disaster or a crisis was in this way shared through the net reaching millions of people, also very distant from the emergency scene, and all without the mediation of journalists. Furthermore, in some circumstances word of mouth news, like that of Twitter, is perceived even more trustworthy than mainstream media [44].

1.1.2 Twitter

Twitter is both a micro-blogging and a social networking site, with around 300 millions active users per month (as reported by Twitter, June 2016 - https://about.twitter.com/company). Twitter's core function allows users to post short messages, or tweets, which are up to 140 characters long. Twitter supports posting of messages in a number of different ways, including through Web services. A remarkable fraction of tweets is posted from mobile devices (around 80%). Messages by a user are displayed as a stream on his Twitter page. Twitter allows a user to follow any number of other users, even if nobody follows him back. It doesn't need a reciprocal connection to be follower of other users. By default, all posted tweets are available to anyone, but users may set privacy preferences in order to make their messages only visible to followers. The three most structural tools of Twitter are: replies, where the use of @ before a username is a way to address a message to that particular user; retweets, messages that are a forward of a content published by another user, typically beginning with RT:@username and hashtags. Hashtags are word prefixed by the symbol # used to tag the conversation topic. The hashtag is the fundamental element of Twitter's user classification system [82], and the platform uses it to aggregate posts according to the topic. Following a particular hashtag, users can monitor conversations about a particular topic or event. This aggregation feature exposes the reader to a variety of tweets, also from users he doesn't follow, entering in contact with diverse information, opinions, or reports on a single topic [74]. Given the shortness nature of tweets, Twitter fits to be used from mobile devices which is especially useful during crisis situations when power supplies may be cut off and short messages are fast and easier to write and send [102]. The brevity of messages in Twitter may be expanded by the introduction of hyperlinks in tweets that lead the reader to other sites or media content for more substantive understanding.

According to current estimates[1], over 350,000 tweets are sent every minute, equating to 500 million tweets per day and around 200 billion tweets per year.

[1] Data accessed on September 30th 2016 on `http:(http://www.guinnessworldrecords.com/news/2016/3/10-years-of-twitter-five-key-tweets-\that-made-record-breaking-history-421461`

In 2016, ten years after the first tweet sent in March 2006, Twitter users are around 300 million. In Italy Twitter users in 2016 were attested around 5 million users.

1.1.3 Twitter, hashtags and emergencies

The amount of information exchanged online can be overwhelming and we may not be able to separate what is relevant from the noise. In Twitter the use of hashtags tends to reduce this effect. Hashtags are words prefixed with the # symbol that works as a message label helping to associate Twitter posts to specific events or discussions. The first use of a hashtag emerged as a tagging convention by Twitter users during an emergency, the California wildfires in 2007. Nonetheless, only a small percentage of Twitter messages contain one or more hashtags [38]. Hashtags are very useful to categorize messages and more organizations are now managing hashtags in a proactive way, identifying specific words or acronyms to label events like a conference, meeting, sports match or TV shows. Hashtags are also important for engagement: periodical recurring hashtags have proven to be the most engaging on average [53].

The use of hashtags may help with the dissemination of information during emergencies but in many cases there is a lack of coordination between government agencies and volunteers in hashtag adoption so information diffusion is not helped [204]. During emergencies a considerable use of hashtags would help users to quickly access information about the event without randomly searching Twitter for relevant messages. Hashtags have also been recognized as a positive factor for retweeting behavior [181]. For a hashtag to go viral its adoption by authority and influencers like media accounts is very important [58].

In recent years codification of Twitter hashtags emerged as an issue in the field of emergency management. The Philippine government was one of the first public agencies to set up a strategy on codified hashtags in 2012 to be used during emergencies [126]. The government declared they had *promoted the use of unified hashtags (#rescuePH and #reliefPH) to monitor, track, and consolidate information before, during, and after a natural disaster strikes.* As reported by Meier [126], the Philippine government suggested a clear convention to create new hashtags by using the local name of the storm in combination with the country acronym PH (e.g. #YolandaPH). An official statement was also distributed to the media and the public to adopt unified hashtags when tweeting about weather-related reports. The Government declared that unified hashtags were very useful in central urban areas where Twitter is more widely used but also in more decentralized ones where Twitter messages conveyed by means of unified hashtags were useful for enforcing communication between government, media and NGOs. The fact that codified hashtags are an important issue in disaster response is confirmed by the publication "Hashtag Standards For Emergencies" issued in 2015 by the UN Office for the Coordination of Humanitarian Affairs, UNOCHA, where standardization of social media hashtags is recognized as a policy that could have major impact on integrating big-crisis data into emergency response. Like the Philippine government strategy, the publication suggests the adoption of a codified syntax to generate new hashtags for next future disasters [138]. In Italy, a proposal for using codified hashtags for weather warnings communication on Twitter was published in 2014. Further details are presented in the next section.

1.1.4 The Italian codified hashtags proposal

During the Genoa flooding of 2014, for the first time in Italy Facebook and Twitter were massively used to organize volunteers and help the city to cope with the disaster. Following the example of Crisis Commons meetings [2], where people discuss the role of technology in humanitarian assistance and disaster relief, in the aftermath of the Genoa flooding (2012/2013) two Crisis Camps were organized in Bologna, Italy, with the aim of opening a discussion with emergency management professionals, volunteers, researchers, public administration communicators and journalists on how to effectively use social media during emergencies. One of the main outputs of the meetings was a proposed *syntax* for codified hashtags to be used in emergencies. In November 2013, the hashtag #allertameteoSAR was proposed by a citizen, an influencer, to coordinate conversation and aids on Twitter during the floods that hit Sardinia. In January 2014 one of the meeting participants (@capitanachab) published the blog post "20 hashtags for a participated civil protection" [3] with a catalog of twenty hashtags to be used during regional weather warning. These are a list of hashtags generated by combining *allertameteo* (weather warning) with the first three letters of the region. The

[2] Further details on `http:\www.crisiscommons.org`

[3] Visit the link `http://capitanachab.tumblr.com/post/74053317969/20-hashtag-per-una-protezione-civile-partecipata`

reason for the regional connotation in the configuration of Italian weather warning systems is that regional meteorological monitoring centers are in charge of local warning for severe weather events. No government recommendations or documents were released to trigger the tactical hashtags adoption during emergencies; it was nothing more than a commitment by some local governments and active citizens to start codified hashtags diffusion. The first regions to actively use them turned out to be those where severe weather events occur most often and flood risk is high.

1.1.5 Promoting the #allertameteoTOS

One of the actions that has been performed within this Ph.D. program was exactly the promotion of the codified hashtag for weather warning in Tuscany. This doctoral thesis is, in fact, the product also of the candidate proper work as communication officer at LaMMA Consortium, the weather service of the Tuscany Region. LaMMA Consortium has built a significant online presence that has its hub in the institutional website counting on over 6 million unique visitors per year. This audience is much broader than Tuscany population, as also assessed in a recent study [80]. Many digital visitors are in fact Italians users, non-resident in Tuscany, and foreigners, consulting weather products and services made available for free on the website, like for instance the outputs of the numerical models for Wind and Sea forecasts over the Mediterranean Sea. In July 2011 with the release of the new web site LaMMA opened an institutional Facebook page and also started using regularly the Twitter account @flash_meteo created in December 2010 to publish weather updates. As happened in many public administrations, the construction of a social media presence for LaMMA Consortium was not a structured process neither was the output of a well-designed communication strategy. It was rather a tolerated initiative proposed by *maverick employees* who were familiar with social networks sites and could foresee the potential benefits for the public administration to jump in [120, 121]. Twitter updates were originally made only by the communication team that slowly started to interest and train forecasters about *publishing basics* of the micro- blogging platform. To gain insights about social media use in case of weather alerts, I attended the Crisis Camp meeting in Bologna on March 2012, where for the first time was discussed the proposal of a standard syntax for emergency hashtags for Italy. Among the participants, motivated by the desire to exploit the potential of new technologies for emergency communications, there were also several employees of public administrations. The meeting led to the birth in April 2012 of the SMEM (Emergency Social Media Management/Messages) Manifesto published on the website socialmediamergencyamanger.com launching the proposal of #smem hashtagging to make emergency tweets recognizable

The SMEM proposal was adopted by LaMMA and promoted through the web site to make people aware of the subject. In the news "Even LaMMA speaks #SMEM: social code for the emergency " is reported that

> *During emergencies we considered the issue of using the correct tag to improve communication. For this reason #SMEM proposal it's not only useful but also essential, in order to have a unique hashtag that can be used and recognized by all.* [4].

Codified hashtags for weather warning were properly used for the first time during Sardinia floods in November 2013, followed, in January 2014, by the blog post of @CapitanAchab (see section 1.1.4). At the end of January 2014, during a red alert for flood risk on the Arno River, the LaMMA Consortium began to use the hashtag #allertameteoTOS to label messages on Twitter dedicated to weather warnings [5] . To promote its use within the organization, we did not follow codified practices but rather a simple illustration to forecasters of the advantages of using this practice on Twitter. The informal process took several months and eventually a working document was internally produced setting out the what, how and when of social media communication in case of weather warnings. A few months afterwards, the codified hashtag for weather alerts entered the check-list of forecasters, as one of the many tasks of the forecasting service shift. It is the first formalization step. In 2015 LaMMA published the Social Media Policy on its website also providing the full list of hashtags used on Twitter, #allertameteoTOS included. Since then, the codified hashtag has been diffused in every news announcing a weather alert on LaMMA website by the following recommendation: "Stay up-to-date by following on Twitter #allertameteoTOS". Since in Tuscany the LaMMA is the point of reference for weather forecasting, the adoption of the codified hashtag in tweets greatly promoted its uptake within citizens, media, municipalities, and non-governmental organizations (NGO)

[4] Visit `http://www.lamma.rete.toscana.it/en/news/anche-lamma-parla-smem-codice-social-l-emergenza`

[5] First adoption of the #allertameteoTOS is reported in a presentation I published on my SlideShare account: `http://www.slideshare.net/ValentinaGrasso?utm_campaign=profiletracking&utm_medium=sssite&utm_source=ssslideview`

working in emergency management.

1.1.6 Twitter as data source

Literature reviews about social media use in disasters [63] attest that Twitter is very used to exchange disaster-related information by all types of users in all kind of disasters. The type of information shared on the microblogging platform include warnings, situational updates, awareness information, but it is also used for personal communication and simple discussions about the events. Facebook, that is the most used social media, is instead preferred for longer messages and an active communication exchange. During disasters Twitter reveals its importance because of some peculiar characteristics. One of them is the hashtags function that allows broadening the number of achievable users well beyond the user's followers [28]. This feature gives Twitter a critical function for citizen's participation during emergencies [63]. Twitter is often used to disseminate crisis messages and access to breaking news. Even if Twitter is not the social media with the greatest numbers of active users (less than Facebook and Instagram), scholars have chosen it as preferred data source for research works on a variety of topics. Among the reasons behind this choice, one of the most important is that tweets are public in the overwhelming majority of profiles, compared to Facebook posts whose visibility is usually limited by strict privacy statements. Furthermore, the Twitter mechanism of following, mentioning and retweeting makes it easier for messages to spread through the networks reaching people that are not "friends" of the user. This makes the platform particularly suitable for information diffusion during critical situations.

Another reason to choose Twitter as data source is that Twitter messages are easily retrievable via the Application Programming Interface (API) put at our disposal by the platform. The access points to Twitter database are Streaming API, for real-time search, or Search API, for historical search. But searching Twitter does not mean getting all of the results. Due to the great volume and frequency of tweets published every day (500 million tweets per day in 2013), Twitter does not have the infrastructure to make them publicly accessible to everyone. Twitter makes available only a small portion of published tweets that for Streaming API is around 1% of Twitter Firehose (up to 10% for Twitter Gardenhose). However, using APIs scholars and analysts have been collecting important amount of messages for research purpose with the help of robust querying systems. To get an idea of what people are talking about and to draw conclusions on communications exchanges, it is enough to have access to statistically significant sample sizes. In any case, we have to be aware that different methods of accessing Twitter data (Search APIs compared to Streaming APIs) can produce quite different sets of results [75]. It could happen that two researchers attempting to gather the same data at the same time will end up with different datasets [30].

Another reason why scholars choose Twitter is because tweets text is short, 140 characters, and thus it is easier to analyze. Twitter is also widely used by media outlets and therefore is suitable for analysis on how information spreads during an emergency.

Many research works have examined the use of Twitter as communication channel across a wide range of applications, from politics [46, 206] to crisis communication in natural disasters [25–27, 40, 93, 140, 180], as discussed in literature review in Chapter 2.

The use of Twitter to create situational awareness during an emergency situation is proven to be a great resource to potentially improve emergency response. This new pattern of communication during disasters has been examined in the past few years by many researchers and emergency management organizations. Many reports and papers have been published in the last 6 years, but few of them are related to the Italian context. Italy has always been particularly vulnerable to natural disasters, even more in the last years as a consequence of the global climate change. Exploring Twitter use during weather-related hazards is important to develop better strategies for effective emergency communication.

1.2 Research design

This research work presents a study of some features of Twitter communication during weather related emergencies in Italy. In particular the research wants to verify if the proposal of using a set of regional codified hashtags to communicate weather warnings on Twitter has been successful in Italy, in order to foster the creation of hashtag-communities during disasters. These communities could be an example of emerging behaviors that during disasters are the expression of a form of "convergence" of individuals and groups committed to help. Social media may lead in fact to group emergence in various forms and functions [122, 176].

To make an assessment we realized a one-year monitoring of the Italian codified hashtags from July 2015 to June 2016. Data retrieval and storage were realized by using the TwitterVigilance platform, developed by the Distributed System and Internet Technologies Lab (DISIT Lab) of the Department of Information Engineering of the University of Florence.

1.2.1 Motivation

This doctoral thesis is strictly linked to my work experience as communication officer at LaMMA Consortium. LaMMA is the Regional body for the Tuscany weather service and is in charge of issuing weather warnings, as part of the Decentralized Functional Center of Tuscany. [6] Over the years, LaMMA built a significant online presence that has its hub in the institutional web site counting on over 6 million unique visitors per year. In 2011 LaMMA entered into social media by opening a Facebook institutional Fan Page and a Twitter account (@flash_meteo) with the aim of enriching the opportunities of interactions with citizens and fostering new practices of public communication and engagement, also in relation to weather warnings communication. My position requires being up-to-date about best practice in emergency communication and social media showed up very quickly to be a great opportunity to deliver weather alerts relying on the fast information propagation they offer. The proper use of hashtags in Twitter during emergency situations turned out to be one of the key issues to solve. Besides the hands-on practice, a great help came from discussions with practitioners and communication officers of public administration and emergency management organizations. The PhD thus was for me an essential part of this professional upgrade directed to improve my understandings and knowledge of social media use in emergency, and potentially transfer it to improve the management of social media profiles of the LaMMA Consortium. The engagement of LaMMA Consortium on this project is demonstrated by a "Research Collaboration Agreement" signed in 2015 by LaMMA Consortium, the Institute of Biometeorology of the National Research Council and the University of Florence, Department of Information Engineering, DISIT Lab [7].

[6] LaMMA is the Environmental Monitoring and Modeling Laboratory for the Sustainable Development. It is a Public Consortium between Tuscany Region and National Research Council of Italy. www.lamma.rete.toscana.it

[7] Distributed System and Internet Technologies Lab of the Department of Information Engineering of University of Florence, directed by Professor P. Nesi, who is my supervisor

Since the publication of the proposal of adopting codified hashtag for weather warnings on Twitter (see description in Chapter 3), I personally and professionally engaged with this project working actively within LaMMA to inform the management and get the approval to start using it in 2014. As outcome, a Social Media Policy and an official hashtags list were formalized. My research is deeply connected to these functional aspects and that's why its main objective was to assess if the proposal of using codified hashtags for weather warnings had any outcomes in Italy.

My research covers an interdisciplinary domain. Scholars from many different research backgrounds investigated social media: from computer science to social sciences and psychology. I relied mainly on my communication studies background, also covering aspects close to disaster sociology, public and emergency communication.

The following sections describe more in detail the research design, main contributions and limits of this research work.

1.2.2 Research questions

As presented in previous sections, main objective of this research is to investigate whether the adoption of a set of regional codified hashtags to communicate weather warnings on Twitter widespread in Italy. The analysis wants to verify if the codified hashtagging in the different contexts led to the creation of hashtag-communities and if the codification of the hashtag helped as well to interconnect emergent behaviors form ordinary citizens with the institutional relief effort. The background of the discussion are the challenges and opportunities made possible by social media in the field of emergency management. The analysis is based on the opportunities offered by Twitter hashtags to create special-purpose virtual community around a topic. By using the hashtag is possible for people around the globe to be part of this virtual community. In this sense the use of social media can help build community disaster resilience.

Main research questions are:

- *In Italy, were the proposed codified hashtags for weather warning employed to communicate on Twitter?*

- *In which contexts did the codified hashtag lead to the creation of online communities?*

- *Which is the role of institutions within these online communities?*
- *Could the use of a codified hashtag help to interconnect official agencies and digital volunteers?*

A one-year length monitoring of Italian codified hashtags was realized, from July 2015 to June 2016. Methodology to analyze data collected is mainly based on similar works on Twitter hashtag data set [22, 29], as explained in Chapter 4.

To compare the diffusion of Twitter codified hashtags in the Twittersphere we used a reference data set of tweets harvested in the same period of time and containing semantically similar but lexically different keywords and/or users related to weather.

1.3 Contributions

Codified hashtagging can be a successful practice to improve communication in weather emergencies, a domain that in Italy has not yet fully been studied. Some scholars in recent years have proposed grammars for hashtags codification during emergencies [111, 178], but this approach has not easily turned into a practice. Many users in fact do not include hashtags in tweets. Moreover, researches attest that the most used tags are often simple and short. This research propose a contribution in this direction, demonstrating that codified hashtags may be successful when sustained by public administrations and local agencies in charge of emergency management. For weather related emergency the adoption by the official weather service and local governments seems to be crucial: if every weather update contains the hashtag it will be easier that media embrace it helping its spread through the network. If this process is not activated, a hashtag will spread only if "Twitter stars" (accounts with many thousands of followers) use it.

This work wants as well to offer a contribution in the research field of disasters studies, in particular focusing on the role of social media for emergent communities. We know by scholars [84, 174] that during disasters people actively collaborate to the relief effort. This is the expression of a form of "convergence" of individuals committed to help. Nowadays, social media provide new forms of engagement and active participation based on information sharing. These emergent groups and behaviors are not easily integrated into a formal emergency response practice, and very often institutional agencies are not prepared to handle it [84, 146]. The use of a codified hashtag may contribute to interconnect formal and informal organizations during the emergency in order to integrate information streams coming from ordinary citizens into institutional emergency management. This analysis explores this hypothesis in the field of weather related emergencies in Italy.

A further contribution of this work is the creation of an ongoing monitoring channel retrieving and storing Italian tweets containing codified hashtags for weather warning. This data set offers an opportunity for scholars and communication officers for further investigations about communication dynamics on Twitter. The data set composed of more than 47.000 tweets (at the day of 26/11/2016), and still growing, represents a *knowledge bank* at disposal of researchers and practitioners to gain insights about the hashtag-communities created around the codified hashtags. The study of communication dynamics of these communities may be helpful to improve communication in similar emergencies.

A last contribution of this work is to have fostered the adoption of Tuscany hashtag for weather warning, #allertameteoTOS, in a public organization formally responsible for warnings issuing. Since January 2014 Consorzio LaMMA started using #allertameteoTOS in warning tweets. Thanks to its broad network of followers (around 14.000), the hashtag gradually entered into use in Tuscany region, being now a *praxis* of Twitter communications during warnings for many Tuscany local municipalities, public institutions, media, emergency management organization and citizens. #allertameteoTOS generated a community of thousands of users, a community that potentially may contribute to make Tuscany more resilient to severe weather events and local emergencies.

1.4 Limits

The work has of course limitations. Main ones are listed below.

- One main concern about Twitter use during emergency is the poor geographic tagging of messages. It's well known that even if it's possible in Twitter to communicate geo-location information by enabling Global Positioning System (GPS) when using the App, in practice only few messages include this information. Scholars indicate that the percentage of tweets that has geo-location meta-data is about 2% [34]. When geo-tagging is not given, it is possible to derive the location of a user by examining his profile description or tweets in his stream; however this derived location is not useful during a crisis because it doesn't reveal the actual location of the person sending the tweets, therefore if this person is a witness reporting on the ground or not. Another way to infer location is to examine tweet content by using specific algorithms based on entity recognition. Given all this, in our work the geo-tagging aspect was considered only by using the analysis of hashtags. Location is in fact perceived as a key information and users tend to use it in an hashtag to highlight its importance. This approach is of course a partial analysis of the geographic dimension within the data set, which could be further investigated.

- Another limitation of the work is that a more comprehensive assessment of the adoption of codified hashtags in tweets during the emergencies would have required an analysis of messages collected by filtering Twitter for other disaster related hashtags (like for Sardinia #alluvioneOlbia or #alluvionesardegna). This would have demanded an adaptive-filtering algorithm on the TwitterVigilance platform (used in this work for message retrieval) to add new keywords to the API querying parameters that is not yet implemented. As an alternative, we confronted the volumes of tweets collected by codified hashtags with tweets related to weather domain, as explained in Chapter 4, about Methodology.

- A qualitative analysis (for example with structured interviews to communication officers of civil protection structures) would have added interesting findings to data discussion, especially for those regions where no adoption emerged. This could be a further analysis to undertake.

- Last but not least, this analysis is bounded in time, so the results are to be considered partial. The diffusion of codified hashtags in different contexts is a dynamic and ongoing process, therefore some of the observations made in this work could be surpassed in the next future by new social or technological scenarios. At the time of this writing (November 2016) a new flooding event in Piedmont has pushed greatly the use of the hashtag #allertameteoPIE [8] that, on the contrary, emerged as one of the least used during the period analyzed in this work.

[8] A new peak was recorded on 25 November 2016 in daily tweets collected by Codified Hashtags monitoring channel in Twitter Vigilance platform (7000 tweets)

1.5 Definitions

We provide a list of most common terms in this domain that are used in this work.

Crowdsourcing. The information obtained by citizens through their collective intelligence and experience is called crowdsourced information.

Disaster. A serious disruption of the functioning of a community or a society causing widespread human, material, economic or environmental losses which exceed the ability of the affected community or society to cope using its own resources [95].

Emergency. A state in which normal procedures are suspended and extra-ordinary measures are taken in order to avert a disaster [147]. Note, for the purpose of this study emergency and disaster are used interchangeably.

Situational Awareness. All knowledge that is accessible and can be integrated into a coherent picture, when required, to assess and cope with a situation [193].

Social Media. Mobile and web-based applications allowing people to communicate and share information across multiple platforms [93].

Twitter. An online news and social networking service where users post and read short 140-character messages called tweets. Registered users can post and read tweets, but those who are unregistered can only read them. Users access Twitter through the website interface, SMS or mobile device app. [202]

Twitterverse. The defined Internet space where all tweets and users exist on Twitter.

About Twitter we define also the following terms:

- Tweet: 140 character publicly viewable message that may additionally contain links and images.
- Twitter Username: (@) a unique Twitter identity.
- Twitter: Real Name The name a user labels himself on Twitter (not necessarily unique).
- Follower: a user that is notified or updated every time the followed user posts a tweet.
- Mention (MT): another Twitter user is mentioned by name in your tweet by using @username.

- Reply: a tweet that begins with another user's @username and it is in reply to one of his tweets.
- Retweet (RT): a tweet that is re-posted by another user by forwarding it. Typically RT begins with the notation RT:@username. Retweets can be unedited RT, a simple forward of the tweet, or edited RT, in this case they do not start with the notation RT:@username.
- Hashtag: a convention used in Twitter to tag certain topics by adding the # symbol before words or text strings (alphabet, numbers or underscore signs).
- Trending topic: A topic that is currently being heavily discussed on Twitter.
- Bot: a Twitter account operated by an automated software application.

Chapter 2

Disasters, society and social media: a literature review

This chapter presents a literature review on disasters, society and social media. A first part of the chapter illustrates works in the domain of sociology of disasters, the theoretical framework used in this work. The following section introduces the topic of social media role during disasters. While reviewing the literature some aspects have been primarily investigated: the reasons for social media adoption by members of the general public; the contribution of social media to situational awareness; social media use by institutions and emergency management organizations. We also present the case of the use of Twitter during Hurricane Sandy in 2012, which is recognized as a turning point in social media use in disasters. Another section briefly presents main metrics and approaches used for social media analysis, including Social Network Analysis. A final section offers a focus on social media and Italian Public Administration.

2.1 A theoretical framework: disasters sociology

In recent years, scholars have turned their attention to the use of social media in disasters, mainly with empirical and best practices studies. Different theoretical approaches have been used to examine social media use in disasters. In this dissertation, we choose as conceptual framework the sociology of disasters, a branch of sociology that is committed to study disasters for their impact on social systems. To better understand social media communication practices and dynamics during disasters, it is important to take into consideration the theoretical acquisitions gained by sociology in this field. We present here a brief overview of the main findings and approaches.

If the focus of sociology is the study of human interactions, sociology of disasters investigates what happens in social systems when a disaster strikes. Even if every disaster is different from the others, it is possible to recognize some elements of commonality and certain behavioral patterns. Since the early research of Prince in 1920 about a ship accident in Halifax Harbor in December 1917, disaster sociologists worked to find and investigate these commonalities [57]. After several isolated researches, it was only with the creation of the Disaster Research Center (DRC) at Ohio State University in 1963, that this field of studies became more structured.

2.1.1 What is a disaster?

The first issue that scholars debated and still debate is the definition of what a disaster is. Sociologists gave many different definitions to this basic question. Following Perry [145], definitions of disasters can be classified into three categories: classic definitions; definitions based on the difference between hazards and disasters; definitions that are socially focused, where disaster is defined by its impacts on a social system.

The first studies about disasters belong to the first approach. They date back to right after World War II and concentrate on the response of the social system to the upheaval. In these approaches, disasters definition relies on the impacting nature of the event. The disaster is seen as a failure of the social system to deliver reasonable conditions of life. Three major definitions emerged in this period. Wallace [195] defines disaster not only in relation to impact but also to the threat of an increase in tensions due to the interruption of normal procedures to reduce those tensions. Killian [108] underline the disasters capacity to disrupt the social order, and people must cope "by departing from the pattern of norm expectations." Moore [133] as well underlines the impact of the disaster in making people adopt new behavior patterns. The disruptive nature of disaster with the consequent loss of lives is a common element of these definitions In 1961 Fritz proposed an alternative definition of disaster more based on a sociological notion of disaster. According to Fritz [69], a disaster is "an event concentrated in time and space, in which a society or one of its subdivisions undergoes physical harm and social disruption, such that all or some essential functions of the society or subdivisions are impaired." ([69, p. 655]) Fritz introduces the notion that a disaster is limited in time and space. Attention is given to the fact that disasters hinder some of the essential functions of the social system. Even more focused on the sociological aspects, Fischer [66] affirms that social scientists are interested in studying social change under disaster conditions.

Stalling [173](1998) definition as well may be included in a classic sociological approach where disasters are seen as disruptions from routines. Others researchers worked on different aspects of this definition. But as summarized by Perry, all these definitions about disasters take into account some basic dimensions: forewarning, the magnitude of impact, the scope of impact, and the duration of impact. Another common feature reported by Perry is the collective trust in the return of the social order, following a typical cycle of stability - disruption - adjustment.

Another approach to defining disasters relies on the distinction between hazard and disasters. This definition is more focused on the nature of the event. Hazard refers to a class of events that can potentially create damages. The classic elaborations of this approach see a disaster as an extreme event that arises when a hazard agent intersects with a vulnerable social system [32, 33]. Close to the traditional view of geographers, Hewitt affirms that disasters are defined by physical agents who have unexpected and unprecedented impacts [85, 86]. Recently, hazards-centered definitions of disaster have moved slightly from an "agent-centered" approach to one more focused on vulnerability. Alexander affirms that when we talk of disasters, we have to include also the social consequences of the events impacting on society [1]. Mileti underlines the social implications of disasters when he says that disasters "are social in nature." ([129, p. 3]) The attention of researchers moves slightly towards the social system instead of considering only the physical event, placing social relationship at the center of disaster studies [156].

The third approach is the one that roots disasters into the social system. Disaster is viewed as a social disruption that begins and ends within the social system. For E. L. Quarantelli, one of the founders of the Disaster Research Center, disasters can be identified by some common features: a sudden onset; the disruption of collective units; the need of actions as adjusting measures; a history unexpected in social space and time; the put in danger of valued social objects [155]. For Quarantelli disasters represent the vulnerability of social structures and systems [156]. Other researchers also focused on the social dimension of disasters. Bates and Peacock [12] view disasters as a process involving the failure of a social system due to its vulnerability. Dombrowsky relates disasters to knowledge and social structure [54]. Some researchers, like Boin [16] and Quarantelli [156], see disasters as a subclass of the larger class of crisis. Barton as well saw disasters as a subclass of a larger one that he labeled collective stress situations. He related disaster to the failure of the social system to maintain expected conditions of life for its members [11].

2.1.2 Disaster types and phases

Another definition issue is related to the classification of different classes of disasters, depending on their causes. In general terms, disasters can be classified into two main groups [151]: 1) Natural disasters, like earthquakes, wildfires, hurricanes, floods, etc.; 2) Man-made disasters, that could be non-intentional disasters (i.e. technological disasters or chemical spills, transportation accidents, etc.) and man-made intentional disasters (i.e. terrorist attacks, etc.).

Researchers have also worked on a definition of different stages that can be recognized during a disaster. Mileti and colleagues in 1975 [130] categorized the four phases of disaster life cycle: response, recovery, mitigation and preparedness. Later, Drabek [56] expanded those findings. Following Drabek (1986) categorization, during a disaster four phases may be recognized. The first is Preparedness that is placed before the impacting event; the second one is Response, as soon as a disaster hits, followed by the Recovery phase. The last phase is Mitigation, actions meant to reduce impacts of disasters on social systems. Each phase may also be divided into two sub-phases: Preparedness (planning and warning), Response (pre- and post-impact), and Recovery (restoration and reconstruction), and mitigation (hazard perceptions and adjustments). In his inventory, Drabek also identified six system levels: Individual, group, organizational, community, society, and international. Even if phases may overlap, this taxonomy has been helpful for many scholars and emergency managers that used it to identify needs, activities, and behaviors peculiar to every phase. Response, recovery, mitigation and preparedness are also the way to define phases of emergency management.

2.1.3 Communication in disasters

The importance of effective communication in emergencies is widely acknowledged. Communication is a core component of disaster preparedness, response, and recovery phases. Effective disaster communication may prevent a disaster or lessen its impact, whereas ineffective disaster communication may lead to even more serious consequences. Disaster communication typically is delivered via the mass media, like television, radio and newspapers, and generally it consists of disaster warning messages and news coverage of the disaster as it unfolds. In their story telling of disasters, mass media shape greatly how the general public perceives and responds to hazards and disasters also influencing individual disaster knowledge, attitudes, and behaviors [88, 158].

There are many aspects of communication during disasters that gained the attention of scholars: the role of the mass media during an emergency; the communication process in which are involved the different components of society; the communication as warning, with studies that analyze how to create effective messages. In his contribution "Unwelcome irritant or useful ally? The mass media in emergencies" [160], Scanlon analyzes the ambivalent relationship between mass media and institutions in emergencies. Mass media are in fact critical for effective emergency communication. They are the link between the general public and the institutions, but their role is not always plain and sometimes they are perceived as a disturbing element. When journalists and reporters crowd in the disaster area, they may create dysfunctions in emergency management and a temporary media black-out is proposed as a way to prevent more people reaching the disaster area and increasing "convergence" [70]. Furthermore, the reporter's need of rapid and updated information is hard to manage by emergency managers or governmental agencies who need to process information before releasing it to the media [160]. If institutions want to make the best use of the media, they have to understand how media work and relate to them with appropriate and trained figures, like a media specialist. The most important function of mass media in emergency is warning. In order to reach the highest number of people, warnings should be released using a mix of communication channels [107].

In emergencies, the main actors of the communication process are citizens, institutions, and the media. The pillars of the communication process between institutions and the general public are trust and credibility, which should be created before the disaster strikes (lombardi2005comunicare).

Broadcasting of warnings is a core function of the communication process in disasters. Disaster warnings usually originate from official government agencies such as national weather services and then disseminated through mass broadcast channels. A warning is a message informing about a danger or a risk and suggesting the desirable protecting behaviors [203]. Scholars have investigated the features a message should have to be an effective warning. Quarantelli assesses that to be effective a warning has to be precise about the threat and about what to do [153]. The message should clearly specify what people should do to protect themselves and to maximize their safety. The language should be plain and should be tailored to communication channels and targets.

A warning is effective if it is conveyed by different channels as to meet people tendency to verify the message through multiple sources. Mileti and Sorensen [131] propose a list of content and style characteristics that a message of warning should have. The topics to consider when writing a warning message are hazard/risk, guidance, location, time, and source. The stylistic aspects are specificity, consistency, accuracy, certainty, and clarity. Similar stylistic features of warning have also been applied to social media messages [184]. Traditional media broadcasting nature is typically a one-to-many process with little opportunity for audience to respond or participate in creating messages. New communication technologies, such as social media, deliver more opportunity for a two-way mediated communication, thus offering the possibility of improved disaster communication, as these technologies have the potential for increased information capacity, dependability, and interactivity [98].

The opportunity of improved and richer disaster communication via social media has captured the attention of communicators and scholars. In the following section 2.2 we will discuss some of the findings offered by a literature review of social media during emergencies.

2.1.4 Organizational aspects during disasters

A field of interest in disaster sociology is how people respond to disasters. Research helped to debunk some myths about people behaviors. Contrary to popular beliefs, Quarantelli's work on the field [154] indicates that during disasters "people as a whole do not panic." Likewise, Fischer's work [65] helped to discredit many stereotypes about people behavior during disasters, which had spread out for a long time also as a consequence of how media represented disaster events. Researches offer a different picture of people reactions to disasters. People tend to become more cohesive, showing pro-social behaviors [150]. When a disaster strikes, affected communities are the first to react to help themselves [197]. Instead of having anti-social behavior, individuals and groups turn out to be mainly altruistic. Organizations and citizens stop their routines and get involved with new tasks and activities related to disaster response. During a disaster citizens are usually the first on the scene. They may be very precious in response and recovery phases, also providing assistance to official agencies. Several studies on the field documented the crucial role played by citizens in emergencies. The research works about emergent groups and organizations are in this area critical.

Several studies took into consideration the phenomenon of *convergence* indicating the informal movement of people and equipment into disaster-affected areas [70]. These movements are called emergent groups or organization. Quarantelli [150] and Dynes [60] proposed taxonomies of organizations active during disasters. It is a scheme focusing on the tasks and structures of groups involved in disasters:

- **Established**: regular tasks / old structures (like the firefighters).
- **Expanding**: regular tasks / new structures (for example volunteers associations or groups whose core activities are non-emergency related).
- **Extending**: non-regular tasks / old structures (a religious group might mobilize its members to deliver food and clothing to the affected population)
- **Emergent**: non-regular tasks / new structures.

Many scholars used this scheme to make hypotheses about organizational behavior in disasters [168, 172].

During a disaster, social organizations change. New behaviors and organizations appear. Emergent groups arise when existing organizations do not meet information needs, resources management and new demands coming out from the disaster [8]. Among the other conditions favorable to emergence is when traditional functions and structures are insufficient to solve the crisis [174] or when the community feels that these structures are inadequate and needs to react [198].

Emergent organizations may include not only people affected by the disaster but also volunteers, emergency workers, government or business organizations [174]. Among the tasks they undertake, there are search and rescue activities, damages assessment, collecting relief supplies, offering emotional support, providing shelter and many other services [55, 174, 197, 198]. The emergence of these independent activities in the affected community is not a rarity. It is instead the expression of a form of convergence of individuals and groups committed to helping.

However, such phenomena may pose problems to official agencies trying to coordinate response activities and to emergency managers during the response phase of a disaster. Too often emergency agencies and governments are not prepared for the response by ordinary citizens during emergencies [84, 146]. Rarely emergency plans take into account the actions undertaken by citizens who are usually not considered in emergency management plans [84]. Emergency staff members see themselves as the only competent responders and push ordinary people aside [84, 146, 161].

Research suggests that citizen convergence during disasters is inevitable, so organizations must find ways to manage the participation of these volunteers who may bring their knowledge, skills, and resources to the rescue effort. Moreover, as Quarantelli [152] affirms, planning should be "with" citizens and not "for" citizens. A bottom up perspective is therefore crucial.

The problem in integrating the citizens' response is related to the fact that in western countries emergency management depends greatly on formal institutional structures based on Command-and-Control approaches [154, 201]. In Command-and-Control approach, the role of emergency organizations is to establish command over chaos and regain control over disorganization. This implies following a bureaucratic hierarchy and structured procedures that left no room for spontaneous volunteerism. Researchers, though, demonstrated that Command-and-Control rarely works well in civilian contexts and a problem-solving approach has to be chosen instead [59, 154, 201]. Management approaches have to rely on an open system that should be decentralized and flexible as to allow a coordination of relief efforts within the different social units [59]. The goal should be problem-solving and not chaos avoiding.

Today with the help of new technologies, scholars recognize that there is also a new form of volunteerism in disasters, that is the activity of digital volunteers. The increasing use of social media and web-based software has allowed citizens to produce and share disaster-related information. In 2010, during the Haiti earthquake a group of students launched one of the first crisis map to monitor damages and urgent needs from the field [125]. Students were crowd-sourcing the information from Twitter and other online sources. Furthermore, an online task force of more than 1000 volunteers from 80 different countries was created to offer translation services and other forms of help. The strength of digital volunteers is that it does not require people to be in the disaster area or to employ a great amount of time. Citizens willing to help may be very far from the emergency area. There is evidence that social media lead to group emergence in various forms and functions, for example coordinating bottom-up relief activities in natural disasters [122,176]. It is a matter of further studies to find ways to interconnect these networks of digital volunteers, local citizens, NGOs and government [122]. In a recent contribution, Boersma and colleagues [15] propose to use a net-centric approach that acknowledges simultaneous citizens participation in several networks, including social media. Incorporating these platforms in disaster response organization is becoming of paramount importance. Social media could also be an effective way to understand how the emergent networks function and interact [15]. Humanitarian agencies are already working to interconnect their formal response procedures with the emergent information streams, coming from web 2.0.

2.2 Social Media and disasters

In recent years, a numerous body of works analyzed how information and communication technologies (ICT) and social media reshaped communication practices during disasters. Social media like Facebook, Twitter, Instagram and YouTube enabled new form of participation during crisis and disasters: new ways of information seeking and sharing and exchanges of assistance. [183] Through social media, citizens may contribute to information availability around crisis by posting an increasing number of photos, videos, maps and reporting. Meanwhile, emergency management organizations are working to understand how to respond to the new content and communication practices made available through these platforms and optimize response efforts with inclusion of data from the people [91,115]. Scholars investigated social media use in disasters by several points of view. Literature is thus very broad and multidisciplinary. In the following literature review, we will focus on some aspects of interest for this research work.

2.2.1 Why people use social media during disasters

While many traditional media remain important disaster communication channels, social media can create opportunities for a two-way dialogue and interaction among organizations, the general public, and individuals [18]. A comprehensive analysis on main reasons behind social media use during emergencies is presented in the report by Frasutino et al. (2012) "Social Media Use during Disasters: A Review of the Knowledge Base and Gaps" [67].

The report summarizes some of the main reasons for SM use during disasters.

- For information needs. People use SM to better understand the risks that they face and reduce uncertainty, especially for real-time updates. During the most recent disasters, Twitter delivered relevant information before than any other traditional media. Sometimes even before it was publicized by the official sources of information.

- For timely information. Social media provide real-time disaster information, which no other media can provide [106, 109]. Social media can become the primary source of time-sensitive disaster information, especially when there is a lack of information from official sources or it is too slow [171]. During the 2007 California wildfires, the public turned to social media because they thought public officials were too slow to provide relevant information about their communities [183].

- To receive unfiltered information. On SM people seek original information or information which has not been revised from traditional media or governments. The public turn to social media because they look for information not previously filtered by any traditional media, organizations, or politicians [118].

- In order to help. During disasters, people often show altruistic behaviors. They want to feel useful in any possible way. Social media offers a simple way of social interaction which in the course of a tragic event helps them to be empathetic with the needs and suffering of other human beings. During disasters, the people turn to social media to organize relief efforts from both near and afar.

- To seek emotional support. During tragic events, people need human contact and emotional support. Nowadays, these can also be found in messages diffused on social media. For those with family or friends directly involved with the disaster, social media can provide a way to ensure safety, offer support, receive status updates [149, 179]. An American Red Cross survey of social media usage was carried out in 2010 it indicated that 24 % of the US population and 31% of the online population would use media to let family and friends know they were safe [49]. Furthermore researchers like Taylor and colleagues (2012) [187] have noted that people caught up in a disaster reported feeling more supported and more optimistic about the future when social media were extensively involved. Gao and colleagues (2012) [71] report how after the 2011 earthquake and tsunami in Japan, the public turned to Twitter, Facebook, Skype and local Japanese social networks to keep in touch with beloved ones while mobile networks were down. Sutton and colleagues (2008) [183] make the point that a disaster is in itself tragic for individuals and communities, so people use social media also because they are in search of human contact and emotional care.

- For humor and levity. Sometimes, sarcastic messages posted on social media may help people to feel positive emotions in difficulties [41,118]. It is important to consider though that humor might not be appropriate in every stage of a disaster. For example, Chew and Eysenbach (2010) found that the public posted fewer humorous comments about the 2009 H1N1 pandemic as the seriousness of the situation increased.

The instantaneous peak in social media use in response to disasters is reported by many academic researches. For example, researchers found that in the half hour before a potential fatal storm hitting a festival in Belgium, SM users published more than 2,000 related tweets [144]. That number increased to more than 80,000 tweets during the first four hours of the disaster. Also, the first reports of the 2008 earthquake in China came from Twitter, not from government [132]

Published literature of social media use in disasters shows that the type of disaster affects the initial activities undertaken by the general public. In natural disasters like hurricanes or floods, people turn to social media in advance to be updated about emergency warnings. In disasters actors often perceive a lack of timely and locally relevant information [89, 165]. During crises, research indicates that SM provide unfiltered, timely, and in-depth communication [101, 149, 186]. Propcopio and Procopio (2007) [149] study about Internet use after hurricane Katrina in New Orleans found that Internet helped to create and maintain social capital, support geographically based communities, activate social networks, and reduce uncertainty. However, social media integrate and not replace traditional media use during disasters. Palen et al. (2010) [140] affirm that during a flooding threat individuals besides looking at social media they also seek for official sources of information such as local news stations. Although social media allow for feedback and two-way communication, often when disasters occur individuals might avoid intervening in discussions and limit to only seek for information. Examining Twitter communication related to the 2011 Las Conchas wildfire threat to Los Alamos National Laboratory, Merrifield and Palenchar (2012) [127] found that several Twitter users were passive, not tweeting or re-tweeting, only searching for information to help mitigate risk and reduce uncertainty.

SM offers the chance to participate in disasters also to people who are physically distant from the event location, like is found in works by Hughes et al. (2008) and Vieweg et al. (2010) [91,193]. The study by Sutton, Palen and Shklovski (2008) [183] demonstrated how social media may also function sometimes as important "backchannels" of communication where citizens may seek and publish information to clarify and integrate information they received from formal and institutional sources.

Another field of social media contribution during disasters is that of collective intelligence, where large groups of people collaborate online to solve complex problems [142]. Torrey and colleagues (2007) [189] found that online means were used to coordinate disaster relief, such as cloths or other items donation. The public may use social media to organize emergency relief efforts. In fact, Starbird and Palen (2010) [176] call *voluntweeters* people who undertake digital and in-person emergency relief efforts. Other researches document the role of Facebook and Twitter for fundraising activities during disasters relief [87]. Another example of collective intelligence is the activity of people using social media to deliver geographically tagged reports, known also as Volunteered Geographic Information (VGI) of crisis events [76, 119]. This geo-located messages can be mapped by volunteers using open source mapping software, like Ushaidi or OpenStreetMap, as it is reported by several studies [77, 210]. Lovari and Comunello (2014) report what happened in Sardinia, Italy, during the floods of 2013, when tweets containing the keywords #allertameteoTOS were used to collect and publish on an Ushaidi map geo-located messages by citizens offering or requesting help thus organizing recovery efforts during and after the floods [45].

2.2.2 Social Media for situational awareness

An interesting point of view in this research is offered by those studies that consider social media's potential contribution to enhance situational awareness [94, 100, 193]. Within the emergency domain, situational awareness describes human perceptions of the different circumstances about a crisis event that allow to read into situations, to make decisions and to predict future outcomes. During a crisis it is crucial for people involved to obtain situational awareness through information that may help to facilitate the understanding of unusually complex situations in order to take better informed decisions. Ireson (2009) in his work on 2007 floods around Sheffield, UK, describes an approach using information extraction, topic and event identification from posts published in a forum that helped to assess how much situational awareness can be raised by the public forum postings. He found that each post can provide some information about the situation, despite the inconsistent quality and conversational nature of the posts. In the work by Vieweg et al. [193], 2010 authors present an in-depth analysis of tweets sent during the 2009 Red River floods and the 2009 Oklahoma City fires, where tweets were collected by filtering tweets published with specific keywords (e.g., #redriver and #okfires). Researchers analyzed tens of thousands of tweets with the aim to identify messages or bit of information that could enhance situational awareness. Authors coded manually tweets as belonging to relevant categories of information as flood-level status, hazard locations, road conditions. Categorization of messages was than helpful for further works using algorithms to extract relevant information. An example is the case of Project EPIC8 (Empowering the Public with Information in Crisis) where researchers developed a natural language processing classifier to help to identify tweets contributing to situational awareness [47, 192]. But generally, given the fragmented nature of Twitter language, automation behind situational awareness derivation is quite difficult to do. Some studies have also investigated how social media data can provide situational awareness for specific crisis-related tasks and domains. For example researchers have developed methods to detect and monitor epidemic by analyzing social media data [39, 51, 136]. Geographic information contained in social media reports have been used in other studies to detect earthquakes or predict earthquake impact and damage [61,157]. In relation to high impacts weather events, the study of Bennett 2013 [13] presents an analysis of traditional and non traditional data sources during hurricane Isaac response in Louisiana in 2012. To have situational awareness for providing timely, life-saving public health and medical response following a hurricane, the study shows that data source like tweets may fill some important gaps providing timely information non produced by traditional media. During a hurricane response where early event detection can save lives and reduce accidents, tweets can provide a source of information for early warning. Another functions of social media during disasters is the contribution in creating a sense of community. People share online their feelings and thoughts, they take care of each other creating a sense of security and community, even when concerns are vast [116, 149].

2.2.3 Social Media use by institutions and emergency managers

As people started to use social media in disasters, formal emergency management organizations started to be pressured to use social media as well. First reason for the adoption is that if members of the general public do not find institutional voices on the social media, they use they will rely for information needs on others sources [179]. Several contributions may be found about this topic. Sarcevic and colleagues (2012) [159] looked at 2010 Haiti earthquake reporting how international groups and medical organizations used Twitter for coordinating activities. Authors assess that even if there is little evidence of a real coordination activity of these groups in Haiti, it is interesting that they used Twitter to be heard by a large audience and thus to use SM as a potential coordination platform. In another work Hughes and colleagues (2014) [92] report of online communication behavior of 840 Fire and Police Departments involved in Hurricane Sandy activity in 2012. Interviews revealed that departments employed social media in different ways, also in creative engagement way. Other researches offer best practice cases to inspire emergency managers on how to use social media. Jin and Liu (2011) [99] analyze how people consume information on social media to understand how the administration should respond during a crisis meeting users expectations. The study of Nilsson (2012) presents e set of indications to be followed by organizations to use social media regarding alerting and warning, from the warning message component, to timing and channels, highlighting the importance of building trust. The work of Temnikova and colleagues (2015) gives some recommendations to government agencies and non-governmental organizations on how to improve readability of tweets, in order to be more clear and easier to understand by the general public [188]. In many cases social media are being introduced within public administrations by digital enthusiasts who work inside the organization. Latonero and Shklovski (2011) [115] investigated the use of social media by the Los Angeles Fire Department (LAFD) in 2009 and revealed that much of LAFD's advanced adoption came from having a single social media "evangelist" in the department. Again in 2009, Hughes and Palen (2012) [90] interviewed 25 Colorado public information officers and reported that even if they wanted to use social media they did not have permission from their management to do so. From interviews it also emerged that they lacked training and resources to assure a social media presence during events. Those who obtained permission to use SM were able only to use it as a one-way communication channel, without any interactivity. In Sutton 2012 [182] authors present the case of ChristChurch earthquake in New Zeland, which shows that coordination and information sharing between volunteers and officials may provide an early solution to the overwhelming need to seek information among the disaster-affected individuals, especially when officials are busy into other life-safety efforts and when human resources are thin. Organizations go through several stages when dealing with social media. At first they start with no social media presence, then they go through a stage of unidirectional use, whether only listening or speaking, only afterwards they enter in a bi-directional phase with positive effects on their activities [50]. In the United States after the turning point of Sandy Hurricane, see next session, the Virtual Social Media Working Group, established by the Department of Homeland and Security, published a report on case studies and well structured guidelines [81] for public organizations on how to use social media during disasters. The report offers precise tips on how to tweet and which is the correct tone of voice to use to inform and reassure citizens.

2.2.4 The case of Hurricane Sandy

Hurricane Sandy represents a great case study to see how social media have been employed during a massive disaster. Hurricane Sandy was a late-season post-tropical cyclone that hit the East Coast of the United States in late October 2012. Because of a full moon, storm tides were 20% higher than normal. As a consequence, Sandy's storm surge was amplified and seawater surged over Lower Manhattan's highways. The water inundated tunnels and subway stations. The electrical system that powers Wall Street went off. Many people died because of the storm impact, thousands were left homeless and millions without power. Property damages were calculated to be around 20 billion dollars. According to the National Weather Service of the United States (NWS), Sandy was the largest Atlantic hurricane on record, as well as the second-costliest Atlantic hurricane in history, only surpassed by Hurricane Katrina in 2005.

Hurricane Sandy was the first New York City social media disaster, in which the most up-to-date stream of information came not from traditional media, like television, radio, but by online media, especially social networking sites.

In a report about Twitter use during Sandy Hurricane, Pew Research Center assessed that more than 20 million tweets were posted on Twitter in a five day period covering the approach and aftermath of Hurricane Sandy in 2012 [1]. The report affirms that in New York usage of Twitter peaked around 9 p.m. on October 29, when the storm hit the metro area. News, information, photos and video made up more than half of all the Twitter conversation. Around 34% of the Twitter discourse about the storm involved news organizations providing content, government sources offering information, people sharing their own eyewitness accounts and still more reposting information created by others. The second largest share of Twitter conversation about the hurricane over these three days, fully 25%, involved people sharing photos and videos. Crucial warnings were sent out through of Twitter and other social media platforms such as Facebook. According to Hootsuite (a social Media Management and Marketing dashboard), #Sandy trended on Twitter while millions of people were without power. Weather forecast sites used Twitter to send crucial weather updates while the hurricane unfolded. Social media provided evidence for the long queues at gas stations and how people were coping without power for days. Power companies used Twitter to inform customers about the many damages and to reassure them with photos and videos of crews working in the field to restore power.

Hurricane Sandy marked a shift in the use of social media in disasters. Throughout Hurricane Sandy, the public turned to social media for updates

[1] Published on Pew Research Center web site: `http://www.pewresearch.org/fact-tank/2013/10/28/twitter-served-as-a-lifeline-of-information-during-hurricane-sandy/`

and assistance, and more than ever before, response agencies, organizations and community groups used social media to organize relief effort and direct resources where needed. A comprehensive report made by Cohen [43] presents a broad overview on social media use during the Sandy Hurricane. We summarize some interesting points. Cohen reports that the New York Office of Emergency Management provided hourly updates and evacuation orders via Twitter. The Governor of New Jersey, Chris Christie, communicated updates about the storm and evacuation orders via his personal Twitter account. People could sign up to receive text alerts from the NY Mayor's Office Twitter account, @nycmayorsoffice, which was used as an alternative source of information to the city's website, especially once people lost power and Internet access. The Federal Emergency Management Agency of the Unites States (FEMA) greatly used social media in addition to traditional ones, to inform and engage the general public across multiple channels for preparedness, when the storm was approaching, and to provide practical and relevant information to those in the affected areas and others outside of the storm's path. FEMA opened a Sandy-specific page on the Fema.gov website to publish all information related to the hurricane, and opened specific Facebook and Twitter profiles as well. As main trusted source of information FEMA attracted millions of users to its online contents. Cohen reports that on Oct. 29, the day Sandy made landfall, FEMA reached more than 300,000 people on Facebook (compared to an average of 12,000 per day), reached 6 million Twitter users with one message (through retweets by individuals and partners), saw 5.800 mentions on Twitter per hour (of the term "FEMA"). Within the huge amount of published messages on line there were also plenty of hoaxes and false images. The verification of information was one of the main task of FEMA online team. Information was verified and rumors were disseminated and dispelled via a variety of tools, including Twitter. FEMA posted popular rumors alongside accurate information in an attempt to dispel inaccuracies and encourage reposting of correct content. A specific website page was set up, Hurricane Sandy: Rumor Control page, which helped to distinguish the truth from false information.

Social media were also the fundamental communication tool for the volunteers of the Occupy Sandy Movement. Occupy Sandy was a grassroots relief effort organized by some of the former protesters involved in the Occupy Wall Street (OWS) Movement. (OWS began in 2011 in New York's Zuccotti Park as an international socio-political movement asking for social and economic justice and new forms of democracy. From New York it spreads rapidly to the world's major cities finding supporters and protesters all over the world). Occupy Sandy emerged to provide mutual aid to communities affected by the hurricane. As the storm hit the city, volunteers turned to social media to organize the relief effort. At first, the initial effort was directed to the OWS members living in the affected areas to make them more prepared to face the storm. There was even an Occupy weatherman sending updates on storm forecasts to all members. Later, the effort spread into local distribution hubs that filtered aid requests and offers in the more affected areas. Occupy Sandy was not an organization but a network of volunteers. Within the first two weeks, Occupy Sandy represented a significant response effort, with an estimated 5,000 to 10,000 volunteers (declared by the Department of Homeland Security of the United States). Its flexible structure was a point of strength to organize a fast response to the Sandy super storm and made it one of the first to move in the most impacted areas. This made Occupy Sandy an important part of the relief effort and in some areas they were the only volunteers doing relief work. An Occupy Sandy website was set up and managed by technical members of the OWS Movement, including members acting from remote places. Facebook and Twitter were heavily employed to coordinate the relief effort in the affected areas when government agencies had not yet reached the most impacted areas. Items and materials needed by storm victims were listed on a wedding registry on Amazon.com allowing people all over the world to donate. Following its principles of mutual aids and people empowerment, Occupy Sandy movement was able to encourage a culture of innovation and fostering micro-local projects, also in the months following the storm.

In 2010, the Department of Homeland Security Science and Technology First Responders Group established a Virtual Social Media Working Group (VSMWG) to discuss the challenges of using social media in public safety. The group collected examples of how social media was used by governmental organizations, volunteers, and ad-hoc groups that were then summarized in a list of best practices for development of all lessons learned. Main findings were published in the report " Lessons Learned: Social Media and Hurricane Sandy" [81].

2.3 Social Media Analysis

2.3.1 Introducing metrics and analytics

Twitter has been the research object of many scholars belonging to a variety of disciplines: computer science, information engineering, linguistics, network science, social sciences, ethnography and so on. We present here a brief overview of the main basic metrics to take into consideration when approaching Twitter data. Bruns and Stieglitz in (2012) [22, 29] present an analysis of the overall metrics of Twitter discussions about different areas of content, in an effort of moving out from analysis of single case studies towards the development of more comprehensive, transferable, and rigorous tools and methods for the study of public communication on Twitter, at different scale. Their study was based on a variety of hashtags data set, collection of tweets obtained by querying the Twitter API for specific hashtags. The point of view of their proposal is, in fact, very much in line with the scope of our research that is also based on hashtags data set. The two scholars affirm that relatively simple metrics, when used comparatively, at scale and over time, can yield analytically productive insights into longstanding questions of media and communication studies:

> *Who are the main actors engaged around a topic or event?*
> *How might we think about the communicative and/or power relations among those actors?*
> *What are the main themes or frames associated with the social media communication around a topic or event?*

Twitter metrics are very much linked to the meta-data and contextual formation that are offered, directly or indirectly, by Twitter entry point: the Twitter Application Programming Interface (API) already offers, directly or indirectly. Table 2.1 presents the list of data and meta-data which the API provide for every single tweet (as appearing in [22]).

Metrics proposed by Bruns and Stieglitz Tweets may also be classified as belonging to one of these category :

- **original tweets**: those tweets which are neither reply nor retweet;
- **retweets**: tweets which contain RT @user; which can be unedited retweets: retweets which start with RT @user; or edited retweets: retweets do not start with RT @user;

Table 2.1: Main meta-data coming with the tweet

text	contents of the tweet itself, in 140 characters or less
to_user_id	numerical ID of the tweet recipient (for @replies)
from_user	screen name of the tweet sender
id	numerical ID of the tweet itself
from_user_id	numerical ID of the tweet sender
iso_language_code	code (e.g. en, fr, ...) of the sender's language
source	client used to tweet (e.g. Web, Tweetdeck, ...)
profile_image_url	URL of the tweet sender's profile picture
geo_type	format of the sender's geographical coordinates
geo_coordinates_0	first element of the geographical coordinates
geo_coordinates_1	second element of the geographical coordinates
created_at	tweet timestamp in human-readable format
time	tweet timestamp as a numerical Unix timestamp

- genuine **replies**: tweets which contain @user, but are not retweets;
- **URL sharing**: tweets which contain URLs.

BASIC METRICS of hashtag data set are the first approach to analyze the collections of tweets. Main metrics proposed are:

- The **number of tweets** in the hashtag data set.
- The **number of unique users** contributing to the hashtag data set.
- The percentage of **original tweets** in the hashtag data set (i.e., tweets that are neither replies nor retweets).
- The percentage of **genuine replies** in the hashtag data set (i.e.,replies that are not retweets).
- The percentage of **retweets** in the hashtag data set.
- The percentage of tweets in the hashtag data set that **contain URLs**.

By comparing many different data sets authors have been able to identify some recurring features connected to some of these metrics. For example, hashtags that show a large percentage of original tweets are usually referred to major media events, like internationally televised entertainment broadcasts (#eurovision, #royalwedding, #oscars) or important sporting events (#tdf, Tour de France) or famous TV shows. On the other hand, hashtags data set showing a substantial amount of retweeting largely fall into a category that they described as *breaking news* which includes also many hashtags related to natural disasters [29].

Authors identify also three main areas of metrics useful to analyze hashtags data sets: metrics that describe the contributions made by specific group of users; metrics that describe activity over time; and a combination of the previous to understand the contributions of specific users over time.

USER METRICS. Users metrics may be divided in Activity metrics and Visibility metrics.

ACTIVITY METRICS are directed at evaluating tweets volumes and users' activity.

- tweets count over time for each user. This is important to highlight those users who are very active compared to the majority who usually only retweets to share information or comments they like;
- original tweets sent, messages which are not replies or retweets;
- sent mentions: tweets which refer to other users. This group is also sub divided in genuine replies and retweets.
- sent retweets, that shold be broken down in edited retweets (tweets that contain RT: @user but do not start with it) and unedited retweets (tweets starting with RT: @user).
- tweets containing URLs to have an indication of the amount of external information single users are introducing in the conversation.

VISIBILITY METRICS are aimed at extracting mentions of users in the data set. Activity metrics are in this case evaluated form the point of view of the users that are the object, not the subject, of the activity. Therefore authors consider the count of total mentions received by a user, the count of retweets and replies received. This approach is helpful to identify different communication styles of users: for example, very often most active users are individuals who tweet and retweet a lot about a certain event, while most visible users are those belonging to news or governmental organizations, which are usually less active but receive many mentions.

Furthermore to differentiate users by their activity rate, Bruns and Stieglitz propose a classification of unique users taking part in the conversation around the hashtag: Lead users: the top 1% of most active users; highly active users: the next 9%; least active users, the remaining 90%. In many cases the first two categories account for the great majority of tweets, while the third one is composed of all users who are part of the long tail of poorly active participants.

Another set of metrics is that related to time: **TEMPORAL METRICS** that are directed to count the number of tweets per period of time. The time unit to consider depends on the typology of event, could be the week, the day or the hour. if time perios is considered, it is also possible to calculate **unique users active over that time period**, distinguishing between original tweets sent by users per period of time and, complementary, retweets sent by users per period of time. All these metrics may be combined in new metrics, depending on the goals of the analysis.

GEOTAGGING CONTENT

Another feature of content that is included in tweets and that may be relevant is whether or not the content of tweet is geo-tagged. It is possible to distinguish two classes of messages: geo-tagged and non geo-tagged. Geo-tagged messages are explicitly associated with meta-data about geographical locations. Non geo-tagged messages do not have this explicit location information, but location may be derived by using implicit textual location index. Twitter, like most social media platforms, allows users to geo-tag the post, hence, this information may be present in the form of meta data. Actually, geo-tagging content depends on several factors. Mainly that the user's device has the ability to mark its location (e.g., via Global Positioning System (GPS)) and that the user has enabled this feature accepting to disclose his location. Basically only a minority of messages includes a machine-readable location information; for instance Burton and colleagues (2012) indicate that this figure is about 2% for Twitter [34]. The location of a user may also be inferred by textual information contained in the tweets text or may be derived by looking at information published in the user's profile or by aggregating information published in his profile. User geocoding refers to the process by which it's possible to derive automatically the location of users when geo-tagging is not active. As for location, it has to be noted that location of a user is less important than the location mentioned by the user in the tweet text.

To geocode a message it's possible and can be done by finding reference of a location named in the text, by means of entity recognition, and assigning geography coordinates to the entity, in order to disambiguate it. This method needs a comprehensive data base of places names, like for instance GeoNames, that is a data base containing 8 millions entries or OpenStreetMap, an open geographical data base. Jurgen and colleagues published a work in 2015 which presents a comparative analysis of geocoding methods [103].

2.3.2 Social Network Analysis

Rapid propagation of information is one of the main characteristics of social media. This propagation process has also been observed during disasters, typically characterized by very high retweeting behavior among users [177]. Graph theory has been very used by researchers to represent propagation dynamics in social networks. Social networks may in fact be represented using graphs where nodes represent users and edges/links represent relations or interactions among those users (for instance following each other, mentioning, retweeting). Social Network Analysis (SNA) provides both a visual and a mathematical analysis of this relationships.

SNA approach has become very popular in the last decades also because it has increased the awareness that we live in a world where social and ecological systems are complex and interdependent. According to Prell [148], Social Network Analysis, as a research filed, emerged in the early Seventies at Harvard University in the Department of Sociology, with Harrison White who was shaping social network concepts with his groups of students. Other researchers, like [68] believe the origins of social network theory should be placed in the work of Jacob Moreno on sociometry, the science of social relations, in the early Thirties [134, 135]. Perhaps the most famous research question connected to social network theory is the "Small World" problem, investigating how long a chain of acquaintanceship would be required to link two people at random in the United States. The problem was solved by Milgrams at the end of the Seventies [190] with the demonstration that two random people in the United States are separated by about six people on average. Though Milgrams never used the expression "six degree of separation", this is how those results became known. At the end of the Nineties, Watts and Strogatz [196] demonstrated that networks from both the natural and man-made world exhibit the small world phenomenon. From social science, social network approach has been applied to a wide range of research fields, from physics to biology, to computer science. Social networking sites' success relies exactly on the application of social networking concepts, linking likely but unknown collaborators using social networking.

SNA is used in social media analytics to model communication patterns and to help identify important people in the network like influential users or opinion leaders, and relevant user communities in social media [180]. It identifies the important users in the network, those who show the strongest influence or act as information bridge between different groups of users. To identify the most important nodes in the network we use Centrality measures. In Social Network Analysis there are different definitions of centrality. Generally centrality measures the importance of a node in the network; but there are several definitions of importance. Centrality measures give us a way to quantify how differently a node can be important.

We give here a short a few **definitions of centrality measures**. Main centrality measures are degree centrality, betweenness, closeness and Eigenvector centrality [162].

- **degree centrality**: in this case the node with most connections is the most important. High value is a measure of prestige, like for a node with many incoming links retweeted by many users;
- **betwenness centrality**: measures the location of the node in the network; high value is for nodes that work as information hubs. Nodes holding high value of betweenness centrality are very important in the network because of their control over information passing;
- **closeness centrality**: closeness considers a node important if it is "close" to, and can quickly communicate with, the others nodes in the network;
- **eigenvector centrality**: a node is important if it is linked to other important nodes. A node with few but important linkers may have a higher eigenvector centrality than one with many links. For example a small twitter account followed by someone with a large audience.

In this research in order to highlight the main factors contributing at the codified hashtag diffusion some **centrality measures** were calculated on the whole period and in peak days. Centrality measures were calculated to identify the influential users in the hashtag-community during selected high-impact events. To calculate centrality metrics and visualize the network graph we used the open-source software Gephi, as further described in section 2.3.3.

2.3.3 Visualizing the network

Graph theory has been used by researchers to represent propagation dynamics in social networks. There are many softwares that can be used for social network analysis. In this work, Social Network Analysis (SNA) was performed with Gephi, an open-source software for network visualization and analysis. We worked on retweets graphs, where each node of the graph is a unique user and each edge (the arch connecting two nodes) is the number of times two nodes retweet each other. Graphs can be directed, where it's possible to identify the source node and the target node, or indirected. For the retweets graph we used a directed graph, where the source node identifies the user retweeting the other, the target node. Statistics of main parameters were calculated with the algorithms provided by Gephi. In our analysis we analyzed two main parameters related to **centrality** measures: betweenness centrality and eigenvector centrality, defined previously in section 2.3.2. In the resulting network graphs, the size of nodes is proportional to centrality measures, betweennees and eigenvector. The network graph is generated by Gephi software using one of the several default Layouts. In this case we used ForceAtlas 2 [97], a force directed layouts algorithm, with default settings. Gephi, in fact, lets the user interact with the representation, manipulate the structures, shapes and colors to visualize data properties and patterns.

To highlight the role played by different categories of users, we attributed a color code to each one: red for Institutions; green for Citizens; orange for emergency NGO; yellow for Media; blue for Weather services. In this way in the network graph it is easy to distinguish more engaged categories of users, by colors, and identify to which category most influential users belong.

2.4 Conclusions

The literature review showed that social media, and particularly Twitter, have been used during emergencies by citizens, news media and institutional agencies as a way to increase information sharing during disasters. The codification of hashtags has been proposed in some cases as a way to improve the coordination of online communication practices among the different social units. As disaster sociology studies assess, online activism during disasters is an example of emerging behaviors where citizens, local and far, affected or not by the event, collaborate actively to the relief effort. Very often official agencies and traditional media are unable to meet very specific information needs on the ground and online communities emerged to answer these needs. In the last years social media played an increasing important role in how people find and exchange information in real time during disasters, in ways that were unimaginable before. Much of this information shared during disasters is produced by online volunteers but rarely emergency plans take into account these information streams [84]. Official agencies are not prepared to integrate and use the information produced by social media. Instead, it is fundamental to acknowledge that virtual communities will be the social reality in the future [152] and societies need to evolve emergency procedures to prevent old and new risks in new ways. Effective emergency planning should be "with" and not just "for" citizens [145]. More inclusive approaches should be preferred to the Command-and-Control approach in emergency management [59, 201].

Citizens' participation through social media is already, and will be in the future, inevitable during disasters, thus it is important that institutional agencies find proper ways to coordinate and manage the participation of volunteers willing to put their knowledge, resources and efforts at disposal of the relief effort. International emergency management agencies, like UNHCR, are already paying attention to procedures and platforms enabling the use of this stream of information. Codified hashtagging could be an effective practice to create a virtual community of users willing to help and share information during emergencies. To establish a codified hashtag for weather emergencies is an example of how to proactively organize and incorporate citizens' participation into a coordinated emergency response. It could represent a solution to make actors and actions of emergent groups more visible to institutional and private members of a community.

In the following chapters (particularly 5, 6, 7, and 8) we will discuss if the proposal of a unique communication tag for weather warnings on Twitter has been applied in Italy. The analysis of the data set of retrieved tweets, filtered per regional codified hashtag, will help to understand whether or not the tag led to an uptake and in which contexts it happened. With the help of social networks analysis, we will also analyze the newborn hashtag-communities to see which were the main users and organizations involved and the role played by institutions and institutional agencies.

We have to consider of course that the codified hashtagging is in many contexts a top-down proposal. This may be considered in contrast to the spontaneity of emergent behaviors. However, it is within such compromise that we can develop effective practices that help to interconnect formal procedures with emergent information streams.

Chapter 3

Italian context: weather warnings, Twitter and hashtags

In this chapter, we describe the Italian weather warning system. The first section presents the organization of public weather service structures, with roles and duties of the network of the Decentralized Functional Centers coordinated by the Central Functional Center of Italian Civil Protection. Moreover, the basic communication procedures employed during weather warnings are presented. The last section illustrates how in Tuscany the codified hashtag was adopted and promoted at the regional level.

3.1 A distributed weather service

Despite what happens in many other European countries, where warnings are issued by National weather services, in Italy a civil National weather service has never been institutionalized. The only National weather services are offered by the Italian Air Force (Aeronautica Militare - AM) which traditionally delivered weather forecast for aviation. In Europe, including Russia and the former countries of the Eastern bloc, there are 31 weather services; only in Italy and Greece the National meteorological services are managed by the military [194]. The postwar Italian democratic governments never considered instituting a national weather service as a civilian public body. After the World war II, meteorological duties from the so called "Ufficio Presagi" (Predictions office) were transferred to the military. This was due not only to the historical role of the Italian Air Force (that even now has competences on weather forecasting for air traffic) but also to the lack of a well-recognized meteorological scientific community within the country [194]. In the early 1990s, the Italian State-Region Conference, a governmental institution that rules the interactions between the central government and the regions, promoted several meetings to establish a national distributed weather service. In May 2001 a governmental decree gave the Italian Department of Civil Protection (DPC henceforth) the task of building a network of Decentralized Functional Centers (DFCs henceforth). The current Italian system for weather warnings has been structured by an Italian Prime Minister directive (on February, 27 2004), that establishes specific procedures for hydrometeorological warnings. The military meteorological service still remains as one of the possible providers of weather data and forecasts. One of the primary functions of the DFCs is to monitor the weather and hydrological situation and identify risks of severe weather, inundations, and landslides. The DFCs are funded by regional governmental administrations and are authorized and coordinated by the DPC.

In 2012 the Italian government proposed the institution by law of the National Distributed Weather Service (although its organization and implementation are still in progress). As explained by Miglietta (2016) [128], the new service is supposed to merge together the activity of the DPC, the Air Force Weather Service and the Regional Hydro-meteorological Services, and should allow for a more rational organization of a system where duties and tasks often overlap. The institution of the National Distributed Weather Service is supposed to overcome the traditional fragmentation of the Italian system. At present, the Italian weather warning system is still managed through the network of the DFCs coordinated by the Central Functional Center of Civil Protection in Rome (CFC henceforth). Next sections briefly describe how the system works and which are the actors involved.

3.2 The Italian weather alerting system

The Department of Civil Protection, lead by the Italian Prime Minister, is the authority responsible for the coordination of policies and activities in the field of civil protection. Established by Law No. 225 of February, 24 1992, it deals at a national level with the prediction, prevention and management of disasters, both natural and human, and of emergency situations. Civil protection, which was originally conceived as a rescue service in emergencies, over the years has become a system for the prediction and prevention of weather phenomena.

The Central Functional Center of the Department of Civil Protection was set up for monitoring and surveillance activities on hydro-meteorology throughout the national territory. The network of DFCs consists of the Central Functional Center, at the Department of Civil Defense, and the Decentralized Functional Centers in the regions and autonomous provinces. To the benefit of the entire network of Functional Centers, the CFC provide to data collection and integration. It also sets up the mosaic of the information generated by existing weather radar systems throughout the country.

Each Decentralized Functional Center conducts forecasting and real-time monitoring of meteorological phenomena with the consequent evaluation of the expected impacts on people and things in a given area. Every DFC, together with the Civil Protection Department and the Regions, contributes to the functioning of the National Warning System. Every functional center must collect and share with the entire network a series of data and information from different technology platforms and a dense network of sensors placed on the national territory. In particular:

- the data obtained from hydrometeorological networks, the national meteorological radar network and the various satellite platforms available for earth observation;
- the spatial data hydrological, geological, geo-morphological and those arising from the landslide monitoring system;
- weather modeling, hydrological, hydro-geological and hydraulic.

Based on these data and modeling, DFCs process the anticipated scenarios, including the impacts on territories. By these assessments, they may also emit bulletins and weather alerts in which are reported both the evolution of the phenomena and the risk levels expected on the territory.

3.3 Weather warning system and communication codes

Warning of the civil protection system, at various territorial levels, is a task and responsibility of the Presidents of Regions and Autonomous Provinces or their representatives (for example, the Director of the Regional Civil Protection). For the purpose of forecasting and the hydro-geological and hydraulic risk prevention, they have divided their territory in early warning areas, geographical areas that are homogeneous with respect to the hydrological and hydraulic effects expected as a result of adverse weather events.

Every day, by 12 AM, the Technical Group for weather forecasting at DPC, produces a weather forecast, valid for the same day of issue and the next. Based on this document, the DFCs prepare the forecast for their own region and carry out an assessment of possible impacts on the ground. If one or more meteorological parameters are expected to exceed specific thresholds of alarm, the Central Functional Center publishes a National meteorological **Watch Bulletin**. This informative document reports the relevant meteorological phenomena forecast for the issue day and the next. Moreover, it gives the expected trend for the following days. The bulletin is published daily at 3 PM on the DPC website.

Each DFC (or the Central Functional Center where a Decentralized Functional Center is missing) makes its own meteorological assessments and deliver them in a Regional Watch Bulletin. If the expected weather events are estimated as particularly severe, the Decentralized Functional Center emits a **Regional Weather Alert**. When the adverse weather conditions may affect more than a region, the Central Functional Center, on the basis of the assessments of decentralized FCs, emits **National Weather Alert** (alerts of severe weather conditions) for the National Service of Civil Protection. In this case, a press release is published on the DPC website.

Typically, weather alerts include: the areas interested by the warning; the duration and timing of the warning; the affected risks (Hydrological or Hydro-geological); the alert level. This last one is expressed through the color code. Warnings are in fact given a color depending on a combination of both the likelihood of the event happening and the impact the conditions may have locally. Alert color could be Yellow, Orange or Red (highest warning level). At central level, the only risks considered for Weather Alert are the Hydrological and Hydro-geological. In some regions, other risks are also considered for the warning assessment. Tuscany for example emits Weather Warnings and Alerts for seven different categories of risk: Hydrological (flooding); Hydro-geological (landslide); Thunderstorms; Gales; Snow; Frost; Sea storm.

The fragmentation of regional Functional Centers has led to differences in the way the weather warnings have been communicated so far in Italy. Only lately, in February 2016, the National Department of Civil Protection officially adopted a working proposal to **standardize the communication of weather warning** throughout the country with **the adoption of the color code**. This happened after eleven years since the Directive of the President of the Council of Ministers has been issued in February 27, 2004, directive that defined the birth of the warning system. This is producing the uniformity both of the warning messages and of the name of the operating steps during emergencies, i.e. preventive measures and emergency management that are activated at the different alerting levels. The goal is to bring the whole Italian Civil Protection system to speak with citizens using the same language, more clearly and comprehensibly than it did up to today. A detail of approved procedures can be found on the web site of Civil Protection Department's website. [1]

Nowadays all Regional governments have to rely on the color code when emitting a weather alert: yellow, orange, or red, from the different scenario forecast and potential impacts on the territories. This implies that every citizen should know that the term yellow alert summarizes an event scenario of flooding of subways, sewer regurgitation, but also critical localized phenomena such as landslides, flash floods, and mudslides. An orange alert means that the forecast events may produce massive floods and landslides in critical geological contexts. A red alert intends that floods and landslides are going to be numerous and of greater magnitude and extent. For each of the three scenarios, there is some danger, from occasional to severe, for the safety of people.

Structures and procedures of the warning system are directed:

- to inform in advance about the possibility that potentially hazardous weather events may occur;
- to activate national/local institutions appointed to verify the capability of emergency management structures to intervene in case of need;
- to put into effect some preventive measures of protection where these are possible, as provided for in the civil protection plans.

Warning communications are in principle conceived to inform the civil protection system. But warnings are also addressed to citizens so that they

[1] Visit the link `http://www.protezionecivile.gov.it/jcms/it/view_prov.wp?contentId=LEG56184`

pay attention to potential risks depending on weather and adopt protective behaviors during emergency situations. Self-protection is in fact the most effective tool to ensure their own safety, especially in case of extreme and sudden events. The alert system is based on forecasts made about 12-24 hours before the events are scheduled to start: so, even if now forecasts are very reliable, they are subject to some level of uncertainty. The alert is designed to achieve the best possible compromise balancing the need to warn in case of hazardous events and to avoid too many false alarms. It is up to citizens to stay always informed about warnings issuing, and to keep track of updates on the appointed institutional channels.

After a weather alert is issued, the communication practices adopted to diffuse the warning depends on the specific contexts. In a case of a multi-regional alert, a press release is published on the web site of the DPC and of regional civil protections. In the last decade, the massive growth of smartphones and use of social networking sites has considerably increased the communication about weather warning and emergencies (see Chapter 2).

In the next section, we will briefly give an overview of social network profiles of the Decentralized Functional Center.

3.3.1 Decentralized Functional Centers and communication scenarios

Despite the fact that an Italian weather warning system exists from 2004, there are still regions today without an independent Functional Center for weather surveillance and monitoring. These Regions rely on the Central Functional Center of the Civil Protection in Rome for weather warning. On this basis, the regional FCs or the appointed structures of Civil Protection make hydrological and hydro-geological assessments and release an alert if needed. Where Decentralized Functional Centers are operative, their institutional configuration may be various. In many regions, DFC is settled within the Regional Agency for the Environmental Protection, the ARPA (Agenzia Regionale per la Protezione Ambientale), which already fulfills many environmental monitoring tasks (like in Emilia Romagna, Piemonte, Fiuli Venezia Giulia, Lombardia, Veneto, Liguria). Somewhere else, it is a service framed within the Department of Civil Protection of the Region or the Province (i.e. Valle d'Aosta; Lazio; Campania). In fewer cases, it is managed by third centers born from the collaboration between authorities and scientific organizations (i.e. Tuscany; Trentino; Abruzzi). It is easy to understand that communication practices are of course very influenced by the different operational contexts and organizations. When the ARPA or a Regional Structure of Civil Protection hosts the DFC, weather warning is just one of the many topics of institutional communication. In these cases, weather warning is just a section of a website or a hashtag within the institutional Twitter account. Among the ARPAs for example, only in Liguria there is a dedicated Twitter account for updates on weather warning. In these cases the press office or communication staff manage the social media account. Since in many public administration organizations, procedures have not yet been updated to the peace of social media applications, it happens that during a weather alert there is no dedicated personnel to manage the Twitter account and communication. Table 3.1 presents the list of Italian DFCs, the related web referral and Twitter accounts if available, and potential hashtags for warning communication. This last information is gained by an analysis of tweets published in each profile and related to weather warning.

As a general framework for the discussion of the use of codified hashtags on Twitter it is crucial to remark some points. Only 5 out of 21 DFCs (regions are 20, but in the case of Trentino there are two separate Functional Centers in charge of warning, one for each Autonomous Province) have no Twitter account. On Twitter, 8 DFC share the account with the agency where they operate within (Region or ARPA); 6 have a Twitter account of their own and 2 have a distinct account for weather alerting posts (Liguria and Abruzzo). See table 3.2. Looking at hashtags, in many cases the tag #meteo is used for weather-related tweets and for warnings. In other two cases, we find a separate hashtag for warning, like #allertameteo or #allertarancio (orange warning). The codified hashtag is used only in Tuscany and Calabria. The Tuscany Twitter account, @flash_meteo, is the one that joined Twitter for a longer time and it also has the largest number of tweets.

3.4 Twittersphere of Public Administrations in Italy

In Italy, social media have a large audience. According to data published by Audiweb [7] in 2015 Italian digital audience was 86.3% of the population, with 41.5 million of Italians who access the Internet from anywhere with a device. In 2015, total digital audience recorded an average value of 28.8 million unique users per month and 21.7 million on an average day. In

Table 3.1: Decentralized Functional Centers: website and Twitter features

Regions	**Web Referral**	**Twitter account**	**hashtags for weather warning**
Val d'Aosta	CFR site	vdaMeteo	
Piemonte	ARPA	ArpaPiemonte	#meteoPiemonte
Lombardia	ARPA	arpalombardia	#meteo
Trentino	CFR	Meteotrentino	
Alto Adige	CFR		
Friuli Venezia Giulia	CFR	meteo_fvg	#meteo #fvg
Veneto	ARPA	arpavstampa	
Emilia Romagna	ARPA	ArpaER	#maltempo
Toscana	CFR	flash_meteo	#allertameteoTOS
Liguria	ARPAL	arpal_rischiome	#rischio#meteo #liguria
Marche	Region	RegioneMarcheIT	
Umbria	CFR		
Lazio	Region	regionelazio	#ProtezioneCivile #RegioneLazio #AllertaMeteo #CodiceGiallo
Molise	Region	ProtCivMolise	#allertameteomolise #AllertaArancione
Abruzzo	CFR	AllarMeteo	
Campania	ARPA		
Puglia			
Basilicata	CFR	pcbasilicata	
Calabria	CFR	Cfm_Arpacal	#allertameteoCAL
Sicilia	Region	DrpcSicilia	#DrpcSicilia
Sardegna	CFR		

Table 3.2: Dedicated Twitter account of DFCs

Twitter account	**Sign-up**	**Followers**	**Tweets**
@vdaMeteo	May-15	219	227
@Meteotrentino	Jul-11	2488	3009
@meteo_fvg	Mar-13	787	3872
@flash_meteo	Dec-10	13600	6998
@arpal_rischiome	Mar-14	2898	1126
@AllarMeteo	Dec-14	144	1475
@Cfm_Arpacal	Jan-14	2531	5704

December 2015 Italians who have accessed the Internet at least once in an average day on computer or mobile device were 22.2 million, 2% more than the previous year; an increase of 7% for mobile Internet use (smartphone / tablet).

Northern areas are those where the percentage of population connected to the Internet is higher (88%), central areas 86%, with lower penetration level in the southern regions (82%). From a demographic perspective it is evident the dissemination among younger segments of the population: 64% of youths between 18-24 years old and 65% of the 25-34 years old are online on an average day. The increase in mobile use also led to the simultaneous growth of the use of social networks. Blogmeter web site reports statistics of June 2015, which attest Facebook as the first social media used by Italian digital audience, with 29 million accounts. Active users per month are 28 million representing approximately 26% of Internet users monthly; about 23 million Italians visit Facebook per day, which is virtually the 100% of the online population. Also according to Audiweb Twitter is the third most used social media, after Facebook and Instagram, monthly users are about 6.4 million (December 2015 data); users are still in sharp decline (-28%) if we consider that in January 2014 they were more than 9 million. A picture of the Italian Twitter-sphere can be found in a blog post published online by Vincos [2] where the author analyzed the relationship between top Twitter users, the ones with at least 100,000 followers (as of August 2013). He recognizes 8 clusters of accounts among the top Twitter users: celebrities (74); media (54); musicians (47); sport athletes (35); brand and Non Gov-

[2] Visit the link `http://vincos.it/the-italian-twittersphere/`

ernmental Organizations (NGO) (21); journalists (18); politicians or political parties (16); fashion brands (8); bloggers (7). Since we are talking about social media, we have of course to consider that the situation is very dynamic, and things change very fast, anyway at the period of the analysis, except for politicians and political parties we find no institutional accounts among Twitter top users.

The spread of social media has also covered the public administration, with a patchy use on the Italian territory. Few studies tried to photograph the use of social media in the Italian Public Administration (PA). Initial reports are those of Arata [2] who investigated time and place of the opening of Facebook and Twitter accounts by the PA in the years 2011-2013. More recently the work of Capineri et al. 2013 [36] presents a report, updated to November 2013, about the diffusion of Italian municipalities on Twitter. At the time of the analysis, there were 461 active official profiles of Italian municipalities on Twitter, about 6% of the total. The region with the highest numbers of active accounts is Tuscany, followed by Valle d'Aosta, Emilia-Romagna, Veneto, and Umbria. The geography of profiles shows the structure of the country, made up of many small and medium-size cities. In the survey, only 1% percent of the profiles belong to the municipalities of large size (with over 600,000 inhabitants), 4% of medium-sized (between 100,000 and 600,000 inhabitants), 44% of low-medium sized municipalities (between 10,000 and 100,000 inhabitants) and 51% are municipalities with less than 10,000 inhabitants. As noted in the study, the reduced size of the municipality does not represent a barrier to the spread of social applications. The diffusion use of Twitter by Italian municipalities began with some pioneering work of institutions such as the city of Rimini in 2007 and then Turin, followed by a regular and gradual growth of others PAs. A section of the study also investigated the most used hashtags as an indicator of the topics discussed on Twitter by institutional accounts. It is interesting to note that among the most discussed topics together with local news, events and public services, it also emerges the weather. Weather forecast became even more important when a warning is issued. Municipalities are in fact on the forefront of the civil protection structures being responsible for the emergency management procedures at municipality level. During emergencies social media have in many cases demonstrated to play a critical role for information diffusion also improving communication among institutions, emergency management organizations and citizens. Social media represented and area of organizational change and redefinition of communication practices within Public Administration [64]. These new practices have been at first related to website and later to social media. It was not a uniform change but more an intermittent process led by early-adopter administrations, which experimented innovative practices acting as a *stimulus* for all the others. In this way public administration started a more participative phase, adopting a new conversational and relational paradigm [120]. However, administrations tend to favor a one-way communication style (a broadcasting style) compared to the dialogic opportunities offered by the social networking platforms [120, 209]. In Italy, most governmental organizations still tend to consider social media alike traditional media, as channels for branding instead of a communication environment for listening and participation practices. Even if social media represent an area of innovation for public communicators, they also highlight the reiteration of old vices and resistances. When communication practices had been reshaped by the Social Media Wave often it happened in a very spontaneous and non-strategic fashion [123]. In a study of Materassi ans Solito (2015) [124] on Tuscany governmental institutions, researchers found out that 53% of the Tuscan municipalities are using web 2.0 channels only for "addressing" population. Facebook is the most widely used platform, followed by Twitter. This last one is employed as publication channel by the press office, or at most for building knowledge and trust networks among organizations, local opinion leaders and the media system, local and national. The opening of the first social profile is driven in most cases by a political will of the Mayor or a Councilor, at the initial moment of its term. This is also in line to what Arata (2013) [2] reports about the national level, where in the most of cases the prevailing organizational structure is the direct management of Facebook account by the local administrator.

A very patchy scenario though, where some administrations are leading the way with innovative initiatives such as the involvement of citizens as digital volunteers, while others are timidly ushering the use of Twitter. Furthermore, not every administration has published a Social Media Policy, an official document where the public administration defines how social media profiles and activities are regulated. During crises, formal protocols and policies are instead essential for effective communication.

In Italy there is not yet an official procedure on how public administrations should communicate on social media during emergencies, except for the "SocialProCiv Maniphesto" which gives some advices for a responsible use of social media during emergencies. The National Department of Civil Protection does not even have a social media profile.

Chapter 4

Methodology

This chapter presents a brief overview of the main analytics to take into consideration when approaching Twitter data. In particular, we examined Twitter analytics that have been used for hashtags data set. Successively, the methodology is presented and discussed. Following sections introduce the retrieval and storage platform used to collect the tweets and the monitoring channels created for this purpose. The last section is designated to discuss basic analytics employed: activity and visibility metrics; coding of users and contents; Social Network Analysis.

4.1 Research Design

The aim of this research project is to verify if, how and where the proposal of using the codified hashtags for weather warning on Twitter has been employed in Italy. Results are in particular interpreted under the theoretical framework of the disaster sociology. Codified hashtag adoption may be interpreted within the context of disaster sociology studies on emergent communities. Hashtag communities arising in a disaster may be considered an example of those emerging behaviors where citizens, local and far, affected or not by the event, collaborate to the relief effort [57, 174]. To establish a codified hashtag for weather emergencies could be an example of how to proactively organize and incorporate citizens' participation into a coordinated emergency response. During emergencies, the lack of communication among institutional agencies and emergent groups is in fact recognized as a problem for emergency management [15, 59, 62, 84, 122]. The analysis of the Italian context aims to verify if the codification of a dedicated regional hashtag for weather warning leads to the creation of online communities and if institutional users are an essential component of these communities. We also investigated if the codified hashtag fostered the communication among digital volunteers and institutions. Main research questions may be summarized as follow:

- *In Italy, were the proposed codified hashtags for weather warning employed to communicate on Twitter?*
- *In which contexts the codified hashtag lead to the creation of online communities?*
- *Which is the role of institutions within these online communities?*
- *Could the use of a codified hashtag help to interconnect official agencies and digital volunteers?*

To perform the analysis, we set up a monitoring project to retrieve and store the tweets containing any of the Italian codified hashtags for weather warning. The online manifesto "20 hashtags for a participated civil protection" published on January 2014 by Twitter user @CapitanAchab on its blog post [1], offers a complete list of the full set of hashtags.

An initial analysis, presented in this work in Chapter 5, was performed by monitoring a set of limited hashtags during a one-month period in 2014. The retrieval was accomplished for 32 days by querying the Twitter API using as keywords three regional codified hashtags. In this first analysis, tweets were retrieved by querying the Twitter API for the selected keywords. The recovered messages were stored in a database. This first case was important to define and test a methodology of analysis to treat the data set that has been validated by the publication of the works in the journal "*Plos. Disasters*" [79]. To organize a one-year long monitoring of the whole set of hashtags, we used a different retrieval method. In fact, due to Twitter API rate limitations, potential network failures, and storage problems, the same procedure seemed too risky. Twitter, in fact, does not give access to the 100% of published messages. The platform makes available to researchers

[1] A participated Civil Protection `http://capitanachab.tumblr.com/post/74053317969/20-hashtag-per-una-protezione-civile-partecipata`

Table 4.1: Italian codified hashtags for weather warning

Italian codified hashtags monitored			
Hashtag	Region	Hashtag	Region
#allertameteoVDA	Valle d'Aosta	#allertameteoUM	Umbria
#allertameteoPIE	Piemonte	#allertameteoLAZ	Lazio
#allertameteoLIG	Liguria	#allertameteoABR	Abruzzo
#allertameteoLOM	Lombardia	#allertameteoMOL	Molise
#allertameteoVEN	Veneto	#allertameteoCAM	Campania
#allertameteoTAA	Trentino A.A.	#allertameteoBAS	Basilicata
#allertameteoFVG	Friuli V.G.	#allertameteoPUG	Puglia
#allertameteoER	Emilia Romagna	#allertameteoCAL	Calabria
#allertameteoTOS	Toscana	#allertameteoSIC	Sicilia
#allertameteoMAR	Marche	#allertameteoSAR	Sardegna

and analysts only a sample of the public data stream. Access is usually provided by means of APIs (Application Programming Interfaces) but only to a sampling data set, which for Twitter is around 1% of the current public data stream [19]. Such limitations may pose problems for tweets retrieval about critical events where millions of messages are published. A continuous monitoring required a more stable system to rely on and also resilient to data peaks or network failure. The solution was offered by TwitterVigilance (TV) platform developed by DISIT Lab of the University of Florence (described in the following section) that was employed in this research for tweets retrieval, processing, indexing, and storage. A set of monitoring channels was created to retrieve and store all tweets containing at least one of the 20 hashtags and other tweets useful to strengthen research assessment. The monitoring period started on July 1st, 2015 and ended on June 30th, 2016, see table 4.1. Furthermore, TwitterVigilance has a number of metrics to assess the efficiency of tweets retrieval at the single channel level. One of these is the estimation of retweets with respect to tweets collected. When a retweet is retrieved, if the reference tweet is missing in the channel the latter is requested and obtained in the 99,5% of times. For small volume channels, this allows to have the 100% of efficiency in recall, and for medium large (over 5 million tweets) the 98% of efficiency.

4.2 The TwitterVigilance Platform

TwitterVigilance is a tool for multi-users collection of tweets and fast statistical analysis (`http://www.disit.org/tv`). It was developed by the Distributed System and Internet Technologies Lab of the Department of Information Engineering of the University of Florence. The platform was partially developed in the framework of an European project as Sii-Mobility Smart City National [2].

4.2.1 Platform main features

Twitter Vigilance architecture is described by the figure 4.1. Twitter provides different ways to access data: REST and Streaming API calls; for all requests, it is necessary to log into Twitter by using OAuth protocol[3]. Every call to APIs returns Twitter data in JSON format. Search API presents a limited number of requests every 15 minutes. The streaming APIs give developers low latency access to Twitter's global stream. Twitter offers different streaming endpoints customized for use type: public, user, and site. Both search and streaming APIs present some limitations regarding the maximum number of tweets per hour, and none of them guarantee that all tweets will be obtained. The TwitterVigilance platform is based on the concept of **"Twitter channel"** defined as a set of simple and complex search queries performed on the Twitter network via crawler. The simplest channel can request tweets referred to a single Twitter user, hashtag or keyword. More complex channels may consist in tens of queries, following the search query syntax of Twitter APIs, obtained by combining keywords, users IDs, hashtags, citations with some operators (e.g., And, Or, From). The Twitter Vigilance is active since April 2015. The TwitterVigilance is able to monitor, follow and analyze slow and fast events on Twitter. A fast event occurs with several hundred thousands or millions of related tweets produced in short time. Slow events may have very few tweets per day or week or their absence for an extended period. The TwitterVigilance collects Twitter data and makes them accessible for the back office processes of statistical analysis, natural language processing (NLP) and sentiment analysis, and for the

[2]For more information on the two projects visit the web site: http://www.sii-mobility.org and http://www.resolute-eu.org

[3]OAuth is an open standard for authorization, commonly used as a way for Internet users to authorize websites or applications, like Twitter, to access their information on other websites but without giving them the passwords.

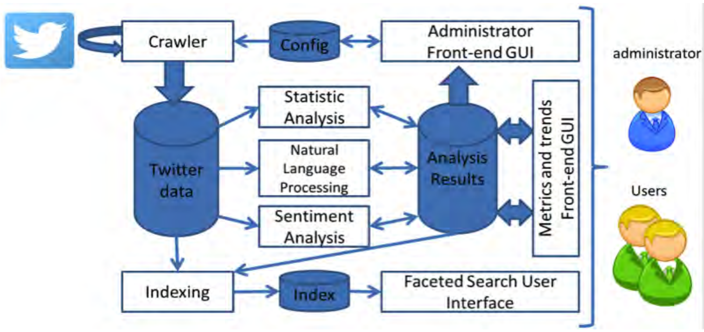


Figure 4.1: TwitterVigilance Architecture

general data indexing, based on NLP on Hadoop [137]. The results of the back office processes are made accessible on a distinct database for the front-end graphic user interface which allows visual analytics, exploration of data results, and the performance of user analysis along time, see for example figure 4.2

4.2.2 Monitoring channels

For the aim of the investigation, multiple channels have been created on TwitterVigilance platform. The main channel is the "Codified hashtags" channel where tweets are retrieved following a multiple queries parameter. All queries are listed in the table 4.1. Other two hashtags were added as well to improve the retrieval: #allertameteoPM, that was used for Piedmont region as alternative to #allertameteoPIE and the tag #allertameteoTS proposed by the city of Trieste (in Friuli Venezia Giulia region) that started a project of digital volunteers trained to correctly publish information during emergencies using the codified hashtag. In addition, other three channels were created with the aim of answering the question about how many users employed the codified hashtags within those usually tweeting about the weather or using Twitter to chat or share information about weather conditions. Below a list and a description of the querying parameters of the three channels:

- "Bad Weather" (Maltempo): is a more general channel created with the aim to collect conversations about bad weather conditions. As

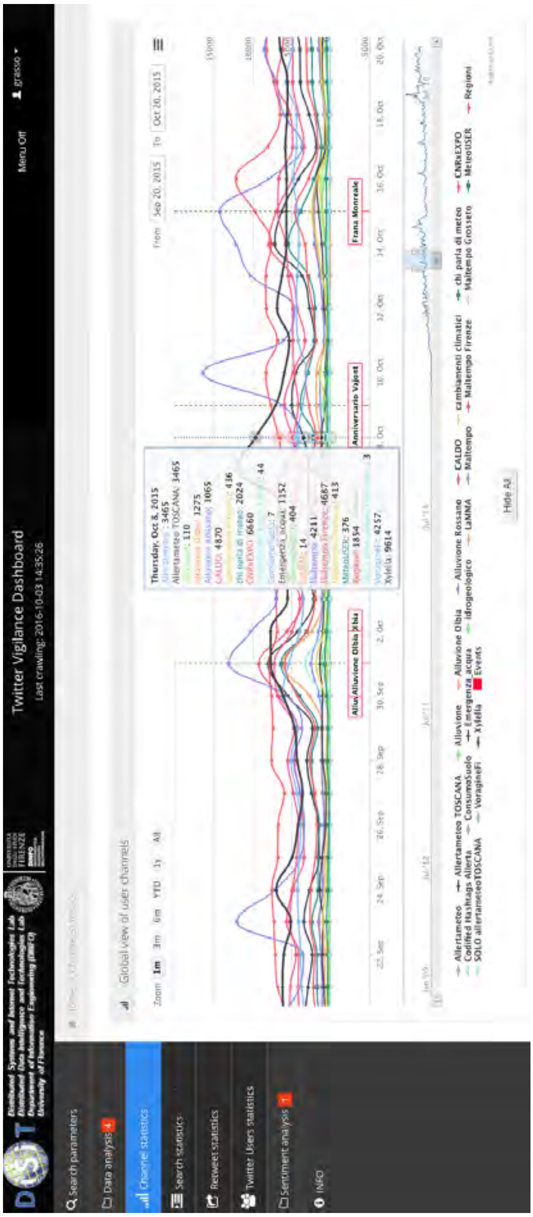


Figure 4.2: TwitterVigilance end-users graphic interface

shown by figure 4.3, querying parameters considered are: #maltempo; #nubifragio; #pioggia; #temporale; allagato OR allagati; bassa pressione; bomba acqua; ciclone; diluvia OR diluvio; grandine; maltempo; nubifragio; piogge; pioggia; piove; piovendo; precipitazioni; temporale; temporali;

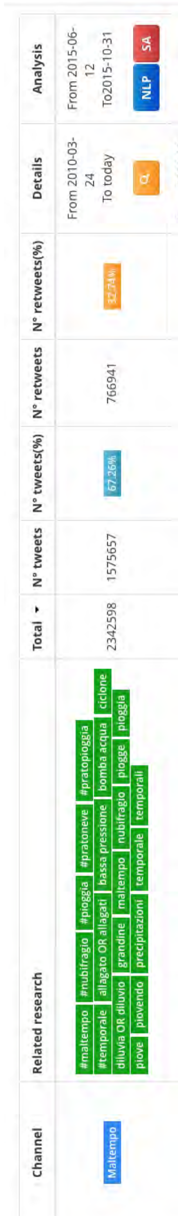
- "Weather" (which talks about the weather): channel composed of all tweets containing the hashtag or simple word "weather".
- "Meteo users" (weather forecasting accounts): a collection of tweets mentioning or coming from well known weather forecasting services (private and public, national and regional) such as: @3Bmeteo; @arpaER; @arpal_meteomare; @centrometeo; @CentroMeteoITA; @flash_meteo; @ilmeteoit;@meteoforumme; @Meteolanterna; @Meteotrentino; @MeteoTweet24; @MeteoWeb_eu; @meteo_fvg; @meteo_toscana; @previsionimeteo; @wwwmeteoit;

As one may imagine, channels having as querying parameters ordinary words are bigger than the codified hashtag channel. The last one is by definition a channel collecting information produced only during high impact events and containing specific tag whereas other channels collect conversations about the weather or about bad weather conditions. Figure 4.4 shows temporal distribution of messages.

4.3 Main analytics

This research has as main driver a communication perspective about the use of Twitter in Italy during particular critical situations like weather adverse conditions and emergencies. Proposed methodology to analyze data set of collected tweets is structured in four main components:

- Twitter activity and visibility metrics;
- a manual coding of users to better understand which category of actors were engaged around the codified hashtags;
- a manual coding of tweet contents to have insights about the information exchanged within codified hashtags conversations;
- a social network analysis to understand information spreading dynamics and central actors.



Channel	Related research	Total	N° tweets	N° tweets(%)	N° retweets	N° retweets(%)	Details	Analysis
Maltempo	#maltempo #nubifragio #pioggia #pratoneve #pratopioggia #temporale allagato OR allagati bassa pressione bomba acqua ciclone diluvia OR diluvio grandine maltempo nubifragio piogge pioggia piove piovendo precipitazioni temporale temporali	2342598	1575657	67.26%	766941	32.74%	From 2010-03-24 To today	From 2015-06-12 To2015-10-31 NLP SA

Figure 4.3: Querying parameters for Maltempo channel

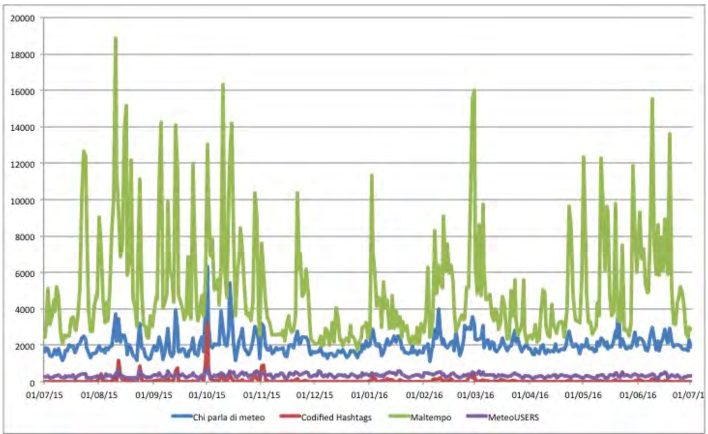


Figure 4.4: Temporal distribution of the four channels

4.3.1 Twitter activity and users metrics

"Codified Hashtags" (CH) data set was analyzed for main metrics: activity pattern over time; volume of different tweets typology over time, differentiating by volumes of original tweets (original messages sent by user) and volume of retweets; volume of mentions and replies; volume of URLs in tweets; combined metrics like ratio native tweets/retweets.
For each data set was also evaluated the number of Active Unique Users, defined as the number of unique users sending original tweets. Visibility metrics were also calculated, in particular: number of favorite tweets; and also most retweeted users and most mentioned users. [30, 31] Relevant metrics of the monitored Twitter channels were performed by using a dedicated R- package developed for the work and released publicly on a Github repository: https://github.com/valenitna/rTwChannel [4].

Main statistics were also made during four selected high impact events occurred within the monitoring period, in order to understand hashtags usage in the different contexts in normal and emergency conditions and get insights on different communication behaviors.

Further details can be found in sections 2.3.1.

[4]R is a language and environment for statistical computing and graphics (https://www.r-project.org/)

4.3.2 Manual annotation of users and content

One of the aims of this work is to describe how codified hashtags have been used in the different contexts, also identifying engaged users and the kind of information exchanged on Twitter. Codified hashtags during emergencies are useful if they may function as useful channels to convey formal and informal sources of information during an emergency. Manual annotation of users and content is also a first step for developing automatic classification algorithms, which needs training data sets.

USERS ANNOTATION: For this purpose, we coded manually the data set of active unique users (users publishing original tweets). The aim was to classify users into main categories and accordingly verify their participation and active role in the conversation around the codified hashtags. To classify users we manually coded the whole set of unique authors by labeling accounts depending on their affiliation, as declared in the profile's description available on Twitter. We considered five classes of unique users as relevant for weather related emergency management. The categories fitting the purposes of this work are: Institutions (governments and public agencies); Media (tv, radio, news and online media); Weather (weather forecasting services or forecasting amateurs associations); Volunteers-NGO (NGOs active in the field of rescue and emergency management); Individuals (accounts of not affiliated individuals; not belonging to any of the above); BOT (computer-generated Twitter profiles that automatically repost certain tweets mentioning a user or a hashtag). By identifying classes of unique users with similar mission and role we compared their communication patterns on Twitter. Comparison of main metrics for each of the monitored codified hashtags was used to assess hashtags adoption in the different regional context and highlight different behavior.

CONTENT ANNOTATION: To identify and to measure the classes of information shared on Twitter through the codified hashtags, content of tweets was manually annotated into defined categories. A set of eleven categories was considered to describe the information communicated in each tweet, paying particular attention to messages contributing to increasing situational awareness (Vieweg, 2010 [193]). Categories considered for coding tweets were: **Advice** (how to cope with the emergency, safety precautions, local emergency numbers to call, advice on how to tweet; websites to follow); **Warning** (tweets about warnings); **Hazard location** (information on hazard localization or reporting about flood or weather impacts on specific locations); **Weather** (information about weather conditions when directly described within the tweet); **Transport Conditions** (updates on road conditions; road closures; airport or public transport malfunctioning); **Evacuation and Closures** (message about closures/opening of public services, schools and scheduled events); **Damage reports** (reported damages on infrastructures or casualties); **Reassurance** (updates of action from first responders and volunteers' activity on the ground); **Resources** (a shared resource, url, picture or video, related to weather or flood update); **Comments** (personal comments, questions, blames) and **News reports** (media resources shared by users). Works by Starbird et al. (2010) [177] and Hughes (2014) [92] guided the identification of categories concerning situational update; two more categories were added to classify media contribution and comments shared by the public broadening the general understanding of emergency impact on the population.

4.3.3 Social Network Analysis

A Social Network Analysis was performed on the data set to explore communication dynamics within the codified hashtags data sets. SNA is used in social media analytics for modeling communication patterns and to help identify important people in the network like influential users or opinion leaders, and relevant user communities in social media [180]. See sections 2.3.2 and 2.3.3 for further details.

To identify the important users in the network, those who show the strongest influence or act as information bridge between different groups of users main centrality measures were applied. In Social Network Analysis, centrality helps to assess the different ways a node can be important.

Centrality measures were calculated to identify the influential users in the hashtag-community during selected high-impact events. This analysis was important to assess if the codified hashtag may foster the communication among digital volunteers and institutional agencies for emergency management. Codified hashtagging could be a proactive way to interconnect official formal response procedures with the emergent information streams coming from web 2.0.

To calculate centrality metrics and visualize the network graph we used the open-source software Gephi, as further described in section 4.4.

4.4 The diffusion of codified hashtags

During the monitoring period, we retrieved all tweets containing the codified hashtags by using the TwitterVigilance platform. We also compared the volumes of retrieved tweets for each one of the twenty codified hashtags to assess which hashtags were mainly used. We considered for the analysis, the contexts where the volume of tweets was high. For the purpose of this work we defined that a hashtag was diffused in a regional context when the total number of users in the channel was above the average value reached within the 20 monitored channels. When a hashtag data set met this condition, the codified hashtag has been considered as sufficiently diffused in the related context.

A further element of the analysis was to try to answer the question of how many people used the codified hashtags respect to potential users. Twitter statistics report an audience of around 6,4 million active users of the social media platform in Italy [5]. Codified hashtags are used in tweets only in case of particular atmospheric conditions leading to warnings or alerts. Therefore to measure the penetration of the hashtag within the *Twittersphere* we compared the tweets collected within Codified Hashtag data set with the volumes of tweets mentioning the accounts related to forecasting services or including lexical keywords semantically related to weather and severe weather conditions. This reference data set was obtained by using the TwitterVigilance platform, where we created three different channels, as described in previous section 4.2.2.

A comparison between the codified hashtags data set and these three data set was made.

4.5 Case studies of high impact events

A more focused analysis was performed on a selection of cases studies related to severe weather events occurred during the monitoring period. These high impact events were identified and discussed more in depth. For the analysis, we selected some severe weather events that affected different areas of Italy to compare hashtags adoption in various contexts. Twitter volume and virality metrics were discussed for the different cases. Social Network Analysis was

[5] Data from Wired, accessed on 29/09/2016, at http://www.wired.it/internet/social-network/2016/04/04/social-media-italia-crollo-twitter-esplode-snapchat/

employed to model interactions between engaged users and to find out the most influential accounts. Since users were coded as belonging to different categories (Institutions, Media, NGO, Citizens and Weather Service), the Social Network Analysis provided a useful visualization to understand which groups played a central role in the different contexts. Three case studies were presented in Chapter 7 covering two main flooding events occurred in Italy in 2015 (Sardinia floods and Calabria floods) and a weather event that impacted many areas of the Italian peninsula, particularly Center-North regions. A further case is discussed in Chapter 8 presenting a very localized and exceptional event taking place in Florence in August 2015.

Chapter 5

A first analysis of codified hashtags

This chapter proposes a first analysis of codified hashtags used for weather warning. The analysis was carried out in 2014, before the creation of the regular monitoring of codified hashtags on TwitterVigilance platform. The analysis investigates the use of the three codified hashtags for Tuscany, Piedmont and Liguria during November 2014, when several flash floods caused great damage in many Italian cities, particularly in Liguria. A first part of the chapter presents the methodology used to retrieve and analyze data. The main metrics and results are presented and discussed. [1]

5.1 Introduction

So far not many researches have investigated the use of social media during weather related emergencies in Italy. Comunello in her book "Social Media e Comunicazone d'emergenza" [45] presents several case studies of social media use during emergencies and in particular one contribution refers to the exceptional flooding in Sardinia in 2013. The study is discussed by Parisi and colleagues, [143] and presents the first analysis of the use of the codified

[1] The core work presented in this chapter has been published as a research article titled "Codified Hashtags for Weather Warning on Twitter: an Italian Case Study" in *PLOS Currents Disasters, 2016* [79].

hashtag #allertameteoSAR used to coordinate information exchange during and post emergency phases. Another notable work is by Giglietto and Lovari (2013) [72] who discuss social media use during an exceptional snowfall in Florence, where authors highlight the active role of Florence municipality's Twitter account @comunefi in establishing the emergency hashtag. The work by Cresci and colleagues (2015) [48] explores the use of social media messages for disaster management, focusing on the automatic detection of critical information for the damage assessment task. Also about automatic detection of valuable information during crisis is the work ny Buscaldi (2015) [35] who applied Sentiment Analysis to detect relevant tweets published during 2014 floods in Genoa.

This section presents a first analysis of the adoption of the weather warning codified hashtags in three different regional contexts.

It is a comparative analysis of the three codified hashtags for weather warning during November 2014 in three Italian regions: Liguria, Tuscany and Piedmont. We analyzed the messages exchanged on Twitter microblogging platform during the selected period of time to identify the amount of information that those hashtags were able to attract. Furthermore, the analysis concerned the characterization of most active users and most used hashtags. The content of messages was also coded to gain insights into the type of information exchanged. In particular content analysis was meant to assess if tweets containing the codified hashtag were more focused on contents contributing to situational awareness, compared to disasters tweets analyzed in similar researches (Starbird 2010, [175]; Starbird and Palen 2010 [177]; Vieweg al. 2010 [193]; Hughes et al. 2014 [92]; Sutton et al 2014 [184]; Bonnan-White et al. 2014 [17]; Soriano et al. 2016 [170]). Findings may be useful to improve information retrieval and processing during disasters, particularly those weather-related [139].

Tweets were collected from November 3rd to December 2nd 2015. We compared three collections of tweets identified by codified hashtags during a period when several severe weather events occurred. The dataset considered is based on the codified hashtags #allertameteoPIE (integrated by #allertameteoPM) (N=2461), #allertameteoTOS (N =3165), #allertameteoLIG (N =29332). Search resulted in 35,558 tweets and 7361 unique tweet authors.

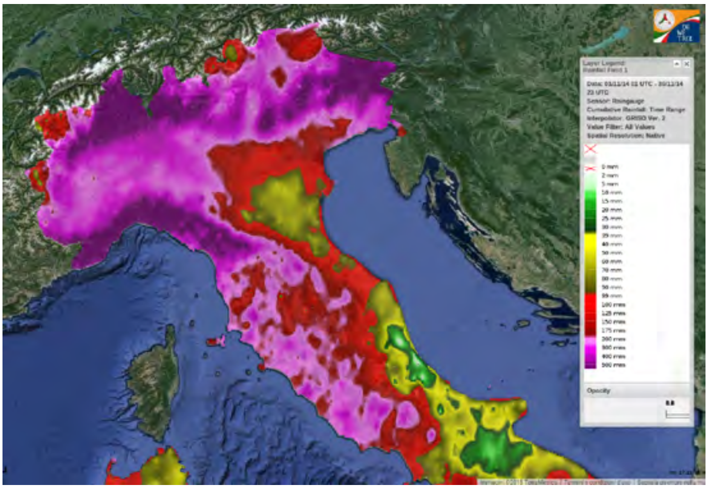


Figure 5.1: Map of cumulated rainfall for November 2014, Dewetra System, Italian Civil Protection Department

5.1.1 Meteorological background

During November 2014 severe weather events (heavy rain and violent thunderstorms activity) occurred in north-western areas of Italy. As well documented on maps of the European Emergency Response Coordination Centre, EU ERCC, (official web site http://erccportal.jrc.ec.europa.eu), in many meteorological national [96] and regional reports [4–6, 112, 113] and in a research paper [166], these events led to flooding in different coastal cities, mainly in Liguria (Genoa), but also in Tuscany (Carrara) and Southern Piedmont. The dates of the three main events were: from 3 to 6 of November 2014; from 9 to 13 of November 2014; from 15 to 19 of November 2014. The corresponding harvesting period of tweets also include another event, less severe and without significant impacts, at the end of November (from November 27 to December 1st). During the period the Italian Civil Protection Department issued 9 red severe weather warnings in Liguria and 3 in Tuscany; 4 orange severe weather warning in Liguria, 5 in Tuscany and 6 in Piedmont (red is the maximum warning level).

The image showed in figure 5.1, provided by DPC Dewetra System (Italian Civil Protection Department) is a map of cumulated rainfall for Novem-

Table 5.1: Channels of codified hashtags

Channels	**Querying parameters**	**Italian region**	**Number of tweets**
#LIG	#allertameteoLIG	Liguria	29932
#TOS	#allertameteoTOS	Tuscany	3165
#PIE	#allertameteoPIE and #allertameteoPM	Piedmont	2461

ber 2014 showing a clear view of weather pressure due to precipitation on the area of investigation.

5.1.2 Dataset and Methodology

Tweets were collected by querying the Twitter Stream API that is the platform application which allows messages and users data harvesting as they are available in the client public timeline. The three regional codified hashtags for weather warning used for Twitter querying were: #allertameteoTOS for Tuscany, #allertameteoLIG for Liguria and #allertameteoPIE integrated by #allertameteoPM for Piedmont. A double querying parameter was used for #PIE because during the considered period both hashtags were proposed on Twitter as codified hashtags for Piedmont region. Approaching the analysis of these three datasets we started considering each collection of tweets as a Twitter *channel* of information. A Twitter *channel* is defined as a time ordered collection of tweets obtained by the Twitter API by using a set of query parameters like hashtags, user accounts or simple word combinations; the time ordered tweets set obtained is named channel and for analogy could be interpreted as an informative content flow semantically linked to related events.

A Twitter *channel* is an efficient way to monitor conversations and facilitate content data mining on defined topics or events identified by the structured combination of search parameters. Relevant metrics of the Twitter channels investigated were performed by using a dedicated R package developed for this work and released publicly on the web.

Querying the Twitter API from November 3rd to December 2nd 2014 for the three codified hashtags a set of 35558 tweets was collected. Three channels were created as shown in Table 5.1.

Through the R-package each channel was analyzed by using the common social-media metrics and also the daily distribution of tweets during the considered period. To give a robust data analysis framework, authors also qualitatively examined and coded the main features of Twitter communication: users, hashtags and content of tweets.

5.1.3 Users annotation

To better describe the channels and identify the communication pattern of different categories of users, a manual annotation of the recovered Twitter accounts was made. The aim was to classify users into main categories and accordingly verify their participation and active role in the codified hashtags. To classify users, we manually coded the whole set of unique authors by labeling accounts depending on their affiliation, as declared in the profiles' description available on Twitter (see 4.3.2).

5.1.4 Hashtags annotation

We also performed a manual annotation of the set of hashtags extracted from each channel. Hashtags coding was intended to identify the popularity of semantically close hashtags used to derive information on which semantic domain users assign relevance to during emergencies. Identifying what kind of information users label as relevant is important for hashtags recommendation during weather warnings.

For each regional channel all hashtags were extracted and manually annotated as belonging to one or more of these classes: **Emergency**, expressions related to emergency warning, damage, severe weather impacts (like floods, mudslide; death reports); **Places**, geographic names like region, county, city, village, suburb, river or point of interest (such as the name of an airport or hospital); **Institution**, government, governors, municipality, mayor, public agencies; **Media**, words identifying media actors or products (TV-shows, TV-news, newspapers); **Weather**, words related to meteorology and forecasting; **Mobility**, words related to street name, motorway, railway and trains; **Social media**, words referring to the social media domain; **Time**, expression of temporal interval; Others, none of the above. Within the class Places was also identified a sub-class named Outside Places (different for all channels), to identify names of places located outside the regional domain of the dominant codified hashtag of the channel. Due to Twitter text limitation and users' shorthand, locations were sometimes expressed in the tweets by means of acronyms, like FI for Florence or GE for Genoa; in this work hashtags expressed by acronyms were not annotated as belonging to the class Places. Because Twitter hashtags cannot include space characters some hashtags turned out to be a combination of two or more words, or even up to a real sentence. In this work composite hashtags were annotated twice, depending on the semantic domain of each lexical component. Because the hashtags annotation was intended mainly to gain a better knowledge about the semantic domain users assign relevance to during emergencies, in Twitter communication, hashtags composed of a combination of words reported a double annotation. For example a hashtag like #alluvionegenova obtained a double annotation, Emergency and Places.

As many studies recognize [23, 117, 205, 208] a hashtag performs multiple functions: it serves as a bookmark of content and a symbol of community membership; in the case of adoption of codified hashtags the community function is supposed to be performed by this whereas the other hashtags included in tweets are supposed to mainly serve as a bookmark of content. As a result, to have a classification of hashtags in categories related to weather, emergency and geographic names it is important to evaluate which other relevant hashtags arose from the channel communities.

5.1.5 Content annotation

We also performed a content analysis of the native tweets of each dataset. Our aim was to identify and to measure the classes of information shared on Twitter through the codified hashtags. Following similar works [177, 184] for content coding were considered only messages from unique authors who posted more than 3 tweets over the considered period. Of the whole 7534 native tweets we coded a sample of 7039 tweets. Only this Refined DataSet (RDS) was considered for message's content analysis. We identified a set of eleven categories to describe the information communicated in each tweet, paying particular attention to those contributing at increasing situational awareness. Further details may be found in section 4.3.2.

Some tweets were coded with more than one category and thus were considered as two messages in the analysis. The reliability of annotations was tested by using Cohen's Kappa [42] calculated comparing message categorizations performed by the two independent panels; an acceptable value was reached with a score of 0.9.

Table 5.2: Features of the three channels (a)

Channel	Total Tweets	Retweets (RT)	Native tweets	TW/RT
#LIG	29932	24017	5915	4
#TOS	3165	2249	916	2.4
#PIE	2461	1758	703	2.5

Table 5.3: Features of the three channels (b)

Channel	Unique Users	Native Tweets Users	Proportion of active Users %	Ratio TW/users
#LIG	5782	1124	19%	5
#TOS	822	171	21%	4
#PIE	757	129	17%	3

5.2 Results and discussion

We began by examining general features of the channels' data set by looking at their composition; we then looked at the daily Twitter activity for each channel; we analyzed participation pattern by comparing the Twitter activity of different classes of annotated users. We examined the hashtags adopted in the three channels and the categories of information posted by the different Twitter users to make an assessment about codified hashtags adoption in these case studies.

The data set is composed of 35,558 total tweets; main metrics are reported in table 5.2 and table 5.3

The vast majority of tweets, 84%, are related to the channel #allertameteoLIG, #allertameteoTOS represents 9% of the whole set and #allertameteoPIE 7%. This is understandable due to the fact that there were several flash floods in Liguria during that period. On November 10th the town of Chiavari was flooded by some minor rivers. On November 15th the city of Genoa (with a population of 880,000 in city and suburbs) was flooded by the river Polcevera and different areas of the city were damaged; it was the second devastating flash flood in Genoa during Autumn 2014, so public attention was already very high.

To describe the channel we considered both the total number of tweets,

Table 5.4: mentions, hashtags, replies and URLs in tweets

Channel	**N. Tweets**	**Native Tweets**	**#**	**Mentioned users**	**URL's**	**Replies**
#LIG	29932	5915	947	1077	2653	395
#TOS	3165	916	277	212	521	22
#PIE	2461	703	218	122	418	24

retweets and replies collected and the subset composed of native tweets, defined as original tweets written by unique Twitter authors. In all three channels the majority of the dataset is made up of retweets, contributing for 70-80% of the total tweets. A high retweeting rate is recognized, in fact, as typical behavior during a disaster event on social media 39 responding to people's need to make information available. The #LIG channel also has the highest ratio between active users and total tweets, with a ratio of 5 (5 retweets per tweet) against 4 and 3 in #TOS and #PIE; the higher participation of #LIG is understandable due to the wider and heavier impacts of severe weather on several cities. One of the metrics considered was the number of unique users participating in the #allertameteoXXX conversation: #LIG presents the highest number with 5782 unique users, compared to 822 in #TOS channel and 757 in #PIE. The majority of these users participated only by retweeting messages; only 20% of users were truly writing messages. This rate is quite similar for all three channels, slightly higher in Tuscany (21%).

Another considered feature was the amount of tweets containing web links (URL's), as external contents or images uploaded on Twitter as links. Researches suggest that the presence of a URL in a tweet is a sign of information richness and is recognized as a key element for situational awareness during a crisis [193]. Almost 50% of native tweets in the 3 channels contained one or more URLs, ranging from 45% in LIG channel to 59% in #PIE channel. This high number is not surprising if we consider that these channels generated by codified hashtags are intended as a way to share information related to warning and emergencies published on the Internet. The ratio Tweets/Retweets and the percentage of tweets with an URL are perfectly in line with crisis event features as suggested by the classification of Bruns (2012) [27]. Tweets including mentions are less numerous than those with URLs: in #LIG and #TOS they reach 28% of native tweets, in #PIE only

13%. Citation habits are discussed in section 5.2.3.

The set of hashtags used in the stream is quite big. It is worth noticing that very few tweets contained only the specific channel codified hashtag: 63 for #LIG, 6 for #TOS and only 2 for #PIE; confirming that hashtagging behavior is very user dependent.

5.2.1 Temporal distribution of channel activity

Daily distribution of tweets gives an idea of the channel activity, that is much related to severe weather events occurring in the period. Looking at the daily distribution of the data set in Figure 5.2 it is quite clear that activity is registered when a warning is issued, peaks are when weather is severe and harmful, and there is no activity when no warning is issued. Tweets collected for each channel correspond with the size and impact of the weather events. The different channels anyhow show some differences. While the #LIG channel reaches highest twitter activity, in general its distribution is quite concentrated on event's peak days with little activity in other periods; #PIE and #TOS channels show a more regular level of activity.

5.2.2 Users participation by categories

To better understand communication pattern of similar Twitter authors, unique users were classified in five categories (Institution, Media, Volunteers-NGO, Weather, Individuals), as earlier mentioned in section 5.1.3.

Table 5.5 shows active users for each category in the three channels and Figure 5.3 shows the contribution of each category to the data set of every channel.

Some differences are easily visible looking at the graph in Figure 5.3. In #LIG channel, Individuals were the most active class, contributing 90% of tweets. The category was as well dominant in #TOS and #PIE. #TOS shows an important contribution of institutional users, the most active category after Individuals, followed by Weather. In #PIE and #LIG the most active class of Twitter accounts, after Individuals. is Media, followed by Weather. In all channels participation of Volunteers is very limited. This could be explained by the fact that during disasters volunteers and NGOs are working hard to face the emergency and probably have no social media staff that regularly post messages.

To gain an enhanced view, we also analyzed the percentage of native

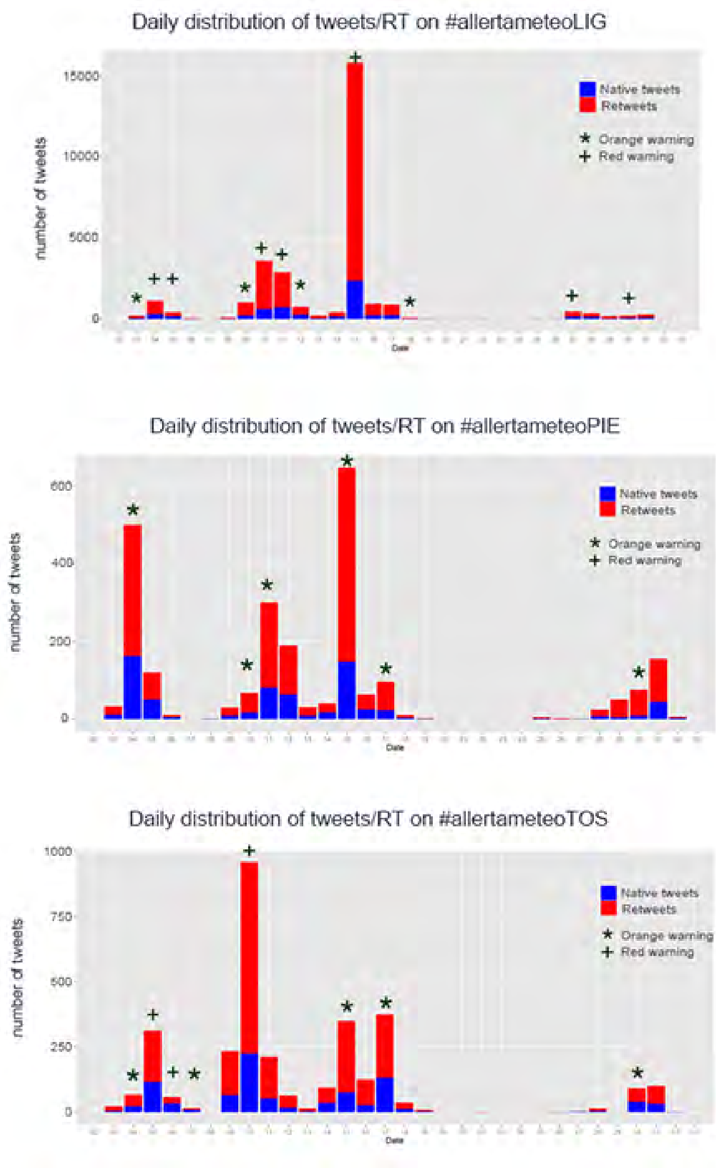


Figure 5.2: Daily distribution of tweets and retweets for each channel and weather warnings issued. Blue bars visualize tweets, red ones retweets. (+) character indicate that a Red warning for severe weather was issued; (*) character that an orange warning was issued.

Table 5.5: Activity rate by category of users in the three channels. (#amXXX stands for #allertameteoXXX)

	#amLIG		#amTOS		#amPIE	
Users category	Tweets and RT	% of native tweets	Tweets and RT	% of native tweets	Tweets and RT	% of native tweets
Institutions	301	47%	703	58%	57	30%
Media	1401	15%	273	49%	468	45%
Volunteers and NGO	227	13%	82	30%	103	23%
Weather	1129	44%	536	19%	282	69%
Individuals	26874	19%	1571	16%	1551	17%

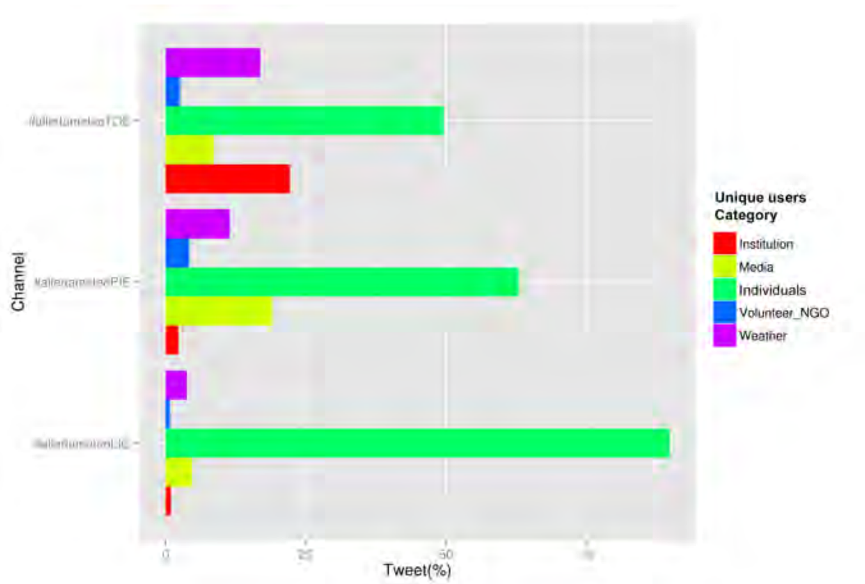


Figure 5.3: Tweets contribution by users category. Bars represent the percentage of tweets published by the corresponding category of unique users, in each channel.

tweets within each category (Table 5.5). This information is interesting to better define the communication pattern of different classes of users and see the difference in the three contexts. In all channels, we noticed that Individuals contribution was mainly expressed in retweeting, with only 16-19% of their messages being native tweets. This is in line with the literature reporting that citizens on social media act as information hub during emergencies [141]. The Institution category, which played a very active role in #TOS channel, shows a different communication pattern in the three channels, with the highest native tweets rate in #TOS (58%) and lowest in #PIE (30%). During disasters institutions are supposed to be the most important and trusted source of information. However, a low level of activity in #LIG and #PIE should not be interpreted as if the institutions were inactive on Twitter, but only that they did not fully adopt the codified hashtag. In Tuscany, the codified hashtag was more supported by the regional weather service, local civil protection offices and municipality, thus gathering more institutional accounts on #TOS channel. In Tuscany the first use of #allertameteoTOS is dated January 31st 2014, by regional weather service Twitter account @flash_meteo.

The Media category made great use of retweets in #LIG compared to #TOS and #PIE, as they frequently retweeted citizens (see following section on citation behavior). Weather category contributed with a high native tweets percentage in #PIE (69% of all tweets were native), in #TOS and #LIG their participation is expressed more in retweeting behavior (81% of Weather tweets were retweets in #TOS and 56% in #LIG). Differences in communication pattern were also due to the more damaging floods impacting the cities of Liguria like Genoa and Chiavari.

The balance between native tweets and retweets may be used to identify the different users' approaches [31]: the *annunciative approach*, posting mainly native tweets, or the *disseminative approach*, posting mainly retweets. Institutions adopted an *annunciative approach* in Tuscany, while in Liguria and Piedmont the approach was more *disseminative*.

5.2.3 Citation behavior: mentioning and retweeting

Another important metric we analyzed to understand communication patterns in the three contexts was citation habits like retweets and mentions. As shown in Figures 5.4 and 5.5, we manually annotated Twitter's users to understand the interactions between relevant categories. The network graphs present mentions and retweets among categories of users in the three channels.

An important metric to understand the channel is to establish which users are the most cited within the hashtag community, measured by the number of @replies and retweets they receive. These are the accounts perceived by the hashtags community as more important and worthy of engaging with [31]. Our analysis proposes this metric at category level, to clarify interactions between the classes of users. Generally, the most active category of users for retweets and mentions is Individuals, with some differences in the three channels. In #LIG, Individuals mentioned and retweeted mainly accounts in the same class. This is also confirmed by the top three mentioned and retweeted authors, who in #LIG channel are individuals. In Liguria, in fact, the members of the public supported the core of the conversation as their tweets amounting to 90% of the whole #LIG data set. Citizens mentioned firstly other citizens, then Weather, Institution, Media and Volunteers-NGO categories. Individuals were the category most cited by all other categories except for Institution. This shows the reticence by a public agency to endorse messages from sources they cannot verify. Institutions tend mainly to cite (mention and retweet) other institutions, in other words sources that are well recognized as trustworthy. The key results for each channel are presented.

In #LIG channel Institution resulted as the category that cited less, a sign of a classic "broadcast" use of social media like Twitter, mainly managed as a medium for publishing information and not as a tool to interact with the public. Media were mainly cited by other media accounts themselves. This tendency to cite accounts belonging to the same class of users also emerges among other categories. The categories act as sub-communities where Twitter users refer to each other. Media and Individuals mentioned Weather accounts that in turn were not so active. They mainly mentioned Individuals and other weather sources.

In #PIE channel, citizens were the most active in mentioning and retweeting. Despite what happened in #LIG, Individuals mentioned mostly Weather accounts and only secondly other citizens, media, and few institutions and NGOs. Weather also turned out to be the most cited class by all the others, also due to a very poor presence of institutions within the codified hashtag community. This inactivity of institutions on the #PIE channel may be explained by the fact that during the severe weather events in Piedmont in October/November 2014 there were two conflicting codified hashtags #aller-

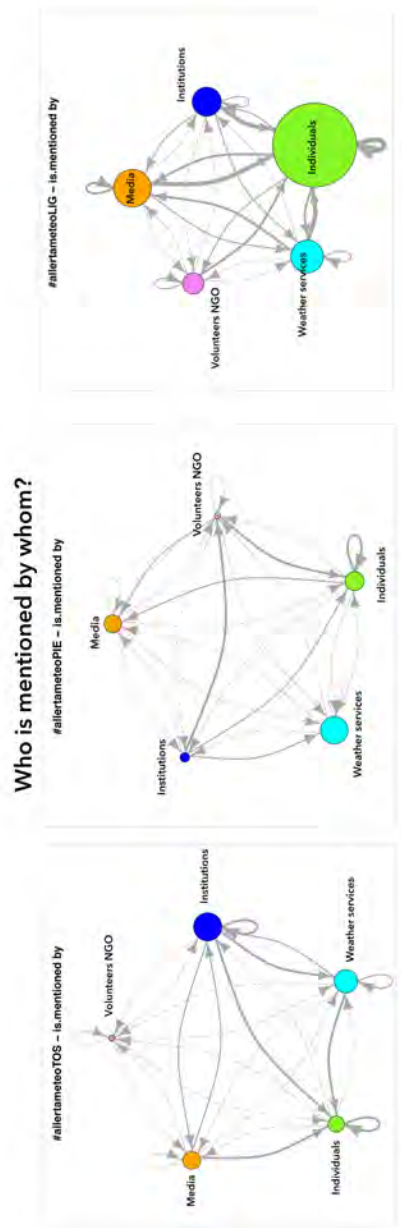


Figure 5.4: Mentions engagement by category of users. The circle dimension represents the number of received mentions; arrows direction indicates the relation is "mentioned by"; the thicker the line is, the higher the mentions received.

Figure 5.5: Retweets by category of users. The circle dimension represents the number of received retweets; arrows direction indicates the relation "is retweeted by "; the thicker the line is, the higher the retweets received.

tameteoPIE and #allertameteoPM (the data set counts 2192 occurrences of #allertameteoPM and 846 of #allertameteoPIE), many institutions also using the latter. Institutions did not cite other institutions. For this reason, the adoption of the codified hashtag in Piedmont was very weak. In #TOS channel, Individuals were very active in mentioning and retweeting citing mostly institutions. #TOS was the only channel where institutions had this active presence within the codified hashtag community. Weather class was the most retweeted by Individuals while Institution was the most mentioned, in a call for action by the public. Institution was the most cited class of users, followed by Weather. It is also interesting to note that institutions cited firstly other institutions and secondly weather's. The #allertameteoTOS promoted by the regional weather service (@flash_meteo) and sustained by municipalities and local civil protection accounts was in this way codified as the official hashtag to be used for weather warning in Tuscany region.

5.2.4 Content analysis

Content analysis was meant to understand the kind of information shared on Twitter during severe weather events. On this purpose, authors coded native tweets in different content categories relevant for situational awareness and analyzed their distribution and activity of different classes of users in the three contexts.

Our analysis showed that codified hashtags were able to aggregate tweets focused on situational updates: messages categories contributing to situational awareness represents around the 85% of the RDS, with minor differences in the three contexts (84% in #PIE and #LIG and 95% in #TOS). This is almost twice the size of the results of Vieweg study [193] on Red River Floods, reporting 49% of on-topic tweets being situational updates. Even in the case of Liguria floods, the messages conveyed through #allertameteoLIG were focused on situational updates. Individuals/citizens have proven to be careful to use the correct hashtag, and many have posted tweets with exhortations to use it strictly for messages concerning emergency management. The most frequently occurring categories of tweets in the codified hashtags RDS were messages of Hazard Location, Weather, Resources, Transport Conditions, Warnings and Advice, as showed in Table 5.6.

An analysis of RDS illustrates its main features and shows similarities and divergences in the three contexts. Around 24% of tweets were coded as Hazard Location. Abundance of Hazard Location messages (situational

Table 5.6: Distribution of contents by categories in the three channels.

	Total	#LIG	#PIE	#TOS
Advice	7%	8%	5%	5%
Comments	12%	14%	4%	2%
Damage/Injury	5%	5%	3%	3%
Evacuation/Closures	6%	5%	4%	11%
Hazard Location	24%	24%	25%	20%
News Report	3%	2%	11%	3%
Reassurance	2%	2%	3%	2%
Resources	15%	17%	12%	10%
Transport Conditions	9%	9%	8%	9%
Warning	6%	5%	3%	13%
Weather	12%	10%	22%	22%

updates of the event containing a geo location reference) confirms that social media have a role in information exchange and that during a disaster individuals provide important updates to complement official information. Hazard location was also the category with the greatest number of retweets.

Messages coded as Weather accounts for 12%, but in #LIG they are half the size (10%) of #TOS and #PIE (22%). A reason could be that in Liguria the regional weather service (ARPAL) was not using the codified hashtag in tweets; Twitter account @ARPAL_meteomare contributed posting automatic updates when a new weather monitoring was issued (on the institutional web site www.arpal.gov.it). Citizens reposted those references by adding the codified hashtags, acting as a hub between trusted sources and the hashtag community. Those kind of tweets were coded as Resources, because the message did not contain any clear textual update but only an external reference. In fact in #LIG dataset Resources category is higher than in #TOS and #PIE. A lot of Individuals published weather related messages but it's worth to notice that writing a textual update within the tweet is by far more useful for people following the Twitter stream than a mere URL sharing that requires the user to open another application to get the information.

About 9% of tweets were coded as Transport Conditions (with similar reach in the three channels), a percentage that is three times the size of what reached by the same category of tweets in Red River Floods case study [193].

In Tuscany this tweets were mostly published by Institutions (78% of Transport Conditions tweets), while in #LIG and #PIE they were mostly coming from individuals (88% and 48%). In Tuscany the existence of an official Twitter account (@muoversintoscana) dedicated to information on the regional road network made it possible to have a continuous and reliable flow of information, also labeled by the hashtag #viabiliTOS.

A 6% of tweets was coded as Warning in the RDS but with interesting differences: 13% of the #TOS channel stream and only 3% of #PIE and 5% of #LIG. A percentage that is in line with previous studies, like the Red River floods by Vieweg (2010) [193], and much higher of the tiny 0,31% reached by warning tweets collected during Yolanda Typhoon in Soriano (2016) [170]. These were the tweets announcing the issuing of a weather warning; the high percentage of tweets in Tuscany it's not a result of higher warnings but rather an outcome of sharing every weather alert on Twitter by the #allertameteoTOS hashtag. Main contribution was by @flash_meteo, the regional weather service, and other Weather accounts: 20% of tweets published by Weather users in #TOS were Warnings, three times as many compared to what happened in #LIG and #PIE (6% and 3% of tweets published by users belonging to Weather related accounts).

Tweets coded as Reassurance were 2% of the dataset, with no quantitative disparity in the three contexts but with different classes of users providing the message. Reassurance is an important class of messages aimed at informing the public that first responders are prepared and active during the event. One would expect a primary role of the institutions called upon to reassure the public that the emergency is under control, but in #LIG dataset these tweets were mainly coming from Volunteers and Individuals (67% and 16% of #LIG Reassurance tweets) rather than Institutions (8%). On the other hand, in #TOS were primarily Institutions to publish Reassurance kind of messages: 35% of tweets coded as Reassurance had institutions as authors, followed by Media users (29%) and individuals' accounts (24%).

Evacuation/Closures messages (updates on evacuation procedures and closures of schools and public offices) accounts for 6% of RDS. They represent important information for the public and are usually issued by institutions and republished by local and national media. In #LIG authors of those messages were mainly individuals (92% of Evacuation tweets), in #PIE Media (45%) and individuals (38%) while in #TOS authors of Evacuation messages have been mainly Media (51%) and institutional (29%) accounts.

A small percentage of tweets were coded as Damages, around 5%. While in Liguria these tweets were around 300 in #TOS and #PIE were barely 25. Individuals were the main authors of tweets coded as Damages in #LIG, confirming the important role of the public as information provider on the ground.

Advice messages amount to 7% of the RDS. These are tweets informing on how to cope with the emergency, giving safety precautions, but also including local numbers to call during emergency and advice on how to properly tweet and use the hashtags. In #LIG 86% of these tweets had Individuals as authors and just 3% Institutions; in #TOS 25% of authors of Advice tweets were Institutional accounts. In #PIE Institution made no Advice tweets at all.

Tweets not contributing to situational awareness, but still on-topic, were coded in two categories, News Report, which consists of tweets about news and media coverage of the emergency; and the category Comments, messages expressing personal opinion, emotions or blames. In #TOS and #PIE this category was scarce (4% and 2% of tweets), while in #LIG Comments were the third most numerous category with 14% of tweets. Because of the damages caused by the Genoa Flood, in Liguria people used Twitter to complain over about government and politicians. Likewise in Tuscany Twitter functioned as a digital *agorÃ* where citizens protested against the mayor and the municipality of the city of Carrara. Criticism was not expressed by the codified hashtag stream though. Other hashtags were used like #alluvioneCarrara (Carrara flood) and #carrarasiribella (Carrara rebels), the last one being a specific tag to organize the public protest. This confirms a more aware use of #allertameteoTOS by the Tuscan community in general.

Compared to similar studies on Twitter usage during flood emergency [110, 193], in the RDS dataset tweets including geo-location information are a higher percentage of all on topic tweets. They are no less than 38% of all tweets (considering that messages coded as Hazard Location, Transport Conditions and Evacuation have always a location reference), that is much more of the 18% reported for Red River study [193] . Tweets that include information about the location of people, the local impact of the hazard, or evacuation sites can improve a better understanding of the situation for individuals reading those messages. Geo-location information is also very useful for the automatic retrieval of relevant information during disasters.

Table 5.7: number of tweets by number of hashtags included in the tweet

Channel	0#	1#	2#	3#	4# or more
#LIG	63	2606	1884	856	506
#TOS	6	241	316	205	148
#PIE	2	219	254	102	126

5.2.5 Hashtags

The last metric analyzed concerned hashtags. The channels considered in this work are generated by three codified hashtags for weather warning, hence every tweet contains at least one hashtag by default. Table 5.7 presents the distribution of tweets by the number of hashtags included in each tweet. It is self-evident that very few users choose to limit hashtagging to only the codified tag. The majority of users added on average one or two more hashtags in the tweet. Around 30% (23% in LIG; 39% in TOS and 32% in PIE) included three or more hashtags, even up to twelve in a single tweet.

The reasons for this behavior could be many. We may suppose that the use of codified hashtagging was at a very early stage in November 2014 (it still is in some Italian regions), thus people choose to combine it with other hashtags emerging spontaneously during the event (like #alluvionegenova or #alluvioneCarrara, included in the top hashtags of the channel). Many users also used more than one codified hashtag in the same tweet to reach a wider public. This was particularly true for tweets containing weather forecasts. We could also argue that, even in the presence of a keyword used to tag a specific conversation and community, as in this case, people tend to introduce more hashtags to highlight information they consider critical inside the tweet, like a place name or a street or even a person. To better understand hashtagging practices, we coded hashtags to gain more knowledge on communication habits. Hashtags included in the channel were annotated in ten categories following the criteria explained in Methodology. The categories were then ranked on the basis of frequency.

As shown in Figure 5.6, in the three channels the majority of hashtags, around 50%, belongs to the class Places. This is in line with the use of Twitter during emergencies to share information on what is happening on the ground, point out problems or report damage, giving specific geographic indications. Tweets containing pictures also tend to include geographic information. The use of geographic information within the codified hashtags

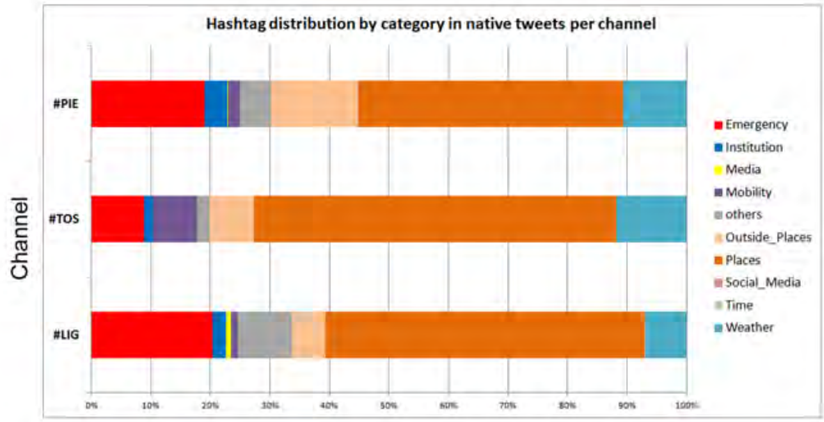


Figure 5.6: Distribution of hashtags following the annotated categories (native tweets only).

community is very important given that a very small percentage of tweets are geo-located and the technique of obtaining location information by querying the users profile is very imprecise. A sub-class Outside_places recorded hashtags referring to a location external to the regional geographic domain of the codified hashtags. References outside of the region are about 10-30% of the class Places. It can be said that tweets were very focused on the local context.

The second most numerous class of hashtags is Emergency, higher in #LIG and #PIE channels (around 20% of total hashtags used) half that in #TOS channel. This is probably due to the severe impacts of bad weather in Liguria, with the flash floods in Genoa and Chiavari. Another point that must be considered is the higher number of users involved in #LIG channel, almost 6000 users compared to less than a thousand in the other two channels. Different users mean different languages and different ways to describe an emergency. Weather was of course another very popular class of hashtags. Weather-related accounts were very active in all three channels and the severe weather was by default the main topic of the codified hashtags community. Hence, it is no surprise that many hashtags belong to the class Weather. The hashtag category related to Mobility is also quite numerous. Indeed, during emergencies users tend to share information about the state of the roads. In Tuscany one of the most active authors was the account @muoversintoscan (Tuscanmobility), which shared a lot information about traffic and transport problems. Media was a well-represented class of hashtags in #LIG but very little used in the other two channels, due to the involvement of PrimoCanale, a local TV station based in Genoa that fully covered the emergency on Twitter by means of its account @primocanale and the account of its main news presenter. Hashtags referred to Institutions were quite represented in #LIG channel, not in #TOS and #PIE, maybe due to the fact that Twitter institutions accounts were not active in the codified hashtag community nor in Twitter, as a result many users (mainly citizens) quoted them in their tweet also in terms of blame. For the Genoa flash flood there were many users criticizing of the Genoa mayor and the Liguria Governor in the media and also on Twitter, and this also emerges in hashtagging practice. Very few hashtags were related to Time category, this kind of information seems not to be considered as a key to become a hashtag.

5.3 Conclusions

This first analysis presents some interesting findings about effective use of codified hashtags on Twitter during weather-related emergencies. It is an exploratory work to investigate the early adoption of the codified hashtags in three different Italian regions (Liguria, Tuscany and Piedmont). Even if the three data sets are quite small compared to those examined in similar works on SM use in natural disasters (#hurricaneSandy, #qldfloods), these results are interesting. The main citation patterns of this data set, like retweeting behavior and number of tweets with URLs, are in line with those reported in the literature for crisis events [27].

The coding of users, hashtags and messages allowed to make evaluations on the flow of information exchanged within the hashtag community and to highlight the role of different users, particularly of institutional accounts that represent main trusted sources. The codified hashtag resulted to be an effective channel of communication. The great majority of information exchanged in this channel related to situational updates, like hazard location, flood updates, damages or injuries reports, road and transports conditions, weather updates, evacuation and closures messages. All these situational updates communicated on Twitter may contribute to situational awareness during an emergency.

Citizens confirmed to be the most active class of users, a peculiarity of crisis events also expressed in the extraordinary retweeting behavior, especially in Liguria, scene of several flash floods during the period considered. Tuscany was the region were Institutions, mainly local civil protection offices, played a central role. In the #allertameteoTOS channel 22% of native tweets were published by Institutions, ten times more than in #allertameteoLIG (only 1%) and in #allertameteoPIE (2%). This active role of the institutions made Tuscany the region where the weather warning codified hashtag gained the widest official agreement. The institutional support is critical for a wider and concrete adoption in an emergency. Piedmont was the region where the use of codified hashtag was weakest, with a lot of tweets posting information about Liguria (the third most used hashtag in the #PIE data set was #allertameteoLIG). The two versions of the regional codified hashtag (#allertameteoPIE and #allertameteoPM) were a sign of this less structured approach.

Hashtags coding revealed that an extraordinary percentage of hashtags were geographic names (50% of items belongs to the category Places), as confirmed as well in the content analysis of tweets. Emergency was the second most represented category of hashtags. In #TOS this category was smaller than in the other two regions, possibly a sign that in Tuscany the semantic related to "Emergency" was felt to have been accomplished by the codified hashtag. Hashtagging practice anyhow reveals a high degree of fragmentation, even higher in the case of a wide participation of members of the public (like in #LIG channel). This makes the use of a codified hashtag even more worthwhile, in order to collect all the information about weather warning under a single umbrella.

Content analysis of tweets showed an almost null rate of off-topic tweets and a very high rate (91%) of tweets related to situational updates. This confirms that the use of a codified hashtag could be very effective in conveying all relevant information during disasters, interconnecting formal and informal sources of information. Compared to similar works, #allertameteoXXX data set count a very high percentage of information related to situational updates, 84% of native tweets, compared for example to 56% reported for Thai floods case study in Konghton 2012 [110], or 61% for Red River floods in Vieweg 2010 [193].

Finally, when considering the possibility to use computer systems to automatically extract information from micro-blogging messages during disasters, as investigated by many researchers, the use of codified hashtags could potentially improve considerably the quality of information extraction due to high rate of on-topic messages and local updates. On this purpose, institutions and emergency managers should consider carefully to support the adoption of the codified hashtags in preparedness activity as a good practice to inform citizens on how to tweet effectively. In Tuscany, where many institutions adopted effectively the #allertameteoTOS, the percentage of tweets focused on situational updates was up to 95%.

Chapter 6

Monitoring of codified hashtags for weather warning

In recent years the use of Twitter codified hashtags for emergency situations has emerged as an important issue. This is confirmed by the publication "Hashtag Standards For Emergencies" issued in 2015 by the United Nations Office for the Coordination of Humanitarian Affairs, UNOCHA, where standardization of social media hashtags is recognized as a policy that can have major impact on integrating big-crisis data into emergency response. A bottom-up proposal to use a set of codified hashtags for weather warning in Italy emerged in 2014. This chapter presents a monitoring of the Italian codified hashtags for weather warning during a one-year period. Particular attention has been paid to regions where hashtags resulted to be more employed, identifying different patterns of communication as emerging from Twitter main analytics and Social Network Analysis. [1]

[1] A preliminary version of the work presented in this chapter has been presented as "Italian codified hashtags for weather warning on Twitter. Who is really using them?" in *16th EMS Annual Meeting and 11th European Conference on Applied Climatology*, held in Trieste, Italy on 12-16 September 2016, and as a short paper published in the Open Access Journal "Advances in Science and Research" - Adv. Sci. Res., 14, 63-69, doi:10.5194/asr-14-63-2017

6.1 Introduction

Many researches have demonstrated the key role played by Twitter during emergencies in the last years; examples are the 2007-2008 California Wildfires [183], Haiti earthquake [167,207], Australian floods [23], Hurricane Ike [93], Hurricane Sandy [92], Red River floods [177,193], Colorado floods [52], Thailand floods [110]. Attending a disaster people with a mobile phone can potentially send information from their location improving situation monitoring; if mobile networks are functioning properly (sometimes they collapse in catastrophic events or are saturated by too many simultaneous users) people may redistribute information acting as a hub for communication diffusion. Through social networking sites, like Twitter or Facebook, the news of a crisis or an accident can be shared by millions of people at a diffusion rate unknown ten years ago. During emergencies a consistent use of hashtags would help users to quickly access information about the event without randomly searching Twitter for relevant messages. In many cases though there is a lack of coordination between government agencies and volunteers in hashtag adoption so information diffusion is not facilitated [204].

In recent years codification of Twitter hashtags emerged as an issue in the field of emergency management, as reported in 1.1.3. The fact that codified hashtags are an important matter in disaster response is confirmed by the UN Office for the Coordination of Humanitarian Affairs (UNOCHA). UNOCHA recognizes that standardization of social media hashtags through a proposed syntax to generate new tags could have major impact on integrating big-crisis data into emergency response [138].

In 2014 an Italian Twitter user proposed a set of codified hashtags to communicate weather warnings in the different Italian regions (see section 1.1.4). One of the main objectives of this research work is to verify if this proposal has been received and adopted by the members of the general public and by institutional organizations. We also want to investigate if the hashtag uptake in some contexts leads to the creation of a hashtag-community, favoring also the communication between institutions and the people. Codified hashtags could be actually a way to overcome the difficulties in integrating the actions of emergent groups and online volunteers with institutional emergency plans [84, 122].

Table 6.1: Italian codified hashtags for weather warning

Italian codified hashtags monitored			
Hashtag	Region	Hashtag	Region
#allertameteoVDA	Valle d'Aosta	#allertameteoUM	Umbria
#allertameteoPIE	Piedmont	#allertameteoLAZ	Lazio
#allertameteoLIG	Liguria	#allertameteoABR	Abruzzo
#allertameteoLOM	Lombardy	#allertameteoMOL	Molise
#allertameteoVEN	Veneto	#allertameteoCAM	Campania
#allertameteoTAA	Trentino A. A.	#allertameteoBAS	Basilicata
#allertameteoFVG	Friuli Venezia G.	#allertameteoPUG	Apulia
#allertameteoER	Emilia Romagna	#allertameteoCAL	Calabria
#allertameteoTOS	Tuscany	#allertameteoSIC	Sicily
#allertameteoMAR	Marche	#allertameteoSAR	Sardinia

6.2 Structure of the analysis

The above-mentioned research objective required to set up a monitoring project to follow and retrieve tweets published using any of the Italian codified hashtags as listed in table 6.1. A specific monitoring channel was created to retrieve and store all tweets containing one of the 20 hashtags from July 1st 2015 to June 30th 2016 and make main analytics for this hashtags data set.

First step was to measure tweets volumes in the given monitoring period and therefore assess which hashtags have been used the most, to successively analyze their adoption more in depth. The Codified Hashtags data set was examined for main metrics: activity pattern over time; volume of original/native tweets (original messages sent by users); volume of retweets; volume of mentions and replies; volume of URLs in tweets; ratio native tweets/retweets; number of active users; visibility metrics concerning number of favorite tweets; most retweeted and mentioned users [30,31]. Top 100 most active authors (users publishing more tweets) were coded. Coding was performed by manually labeling accounts depending on their affiliation, as declared in users' profiles description available on Twitter. Category considered for coding are: Institutions, Media, Citizens, Weather (services), NGO and BOT, as fully described in section 4.3.2. For data sets of frequently adopted codified hashtags, we performed a comparison of main metrics to assess hashtags adoption in the different regional contexts and highlight different communication behaviors.

A final section is directed to compare the usage of codified hashtags in the Italian Twittersphere with tweets generically related to bad weather conditions.

6.3 Main results

In this section, main results of the codified hashtags monitoring are presented. At first we present the main features of the data set and a comparison between hashtags. Some definitions of terms used during the analysis are provided:

- Native/original tweets: tweets originally produced by users;
- Retweets: tweets that are a repetition on another user's update, usually beginning with RT;
- Full tweets: the whole set of tweets and retweets of the specified data set;
- Active users: unique authors publishing native tweets and not just retweeting;
- Active period: number of days where at least one tweet containing the codified hashtag was published;
- Activity rate: percentage of the active period in relation to the whole considered monitoring period.

6.3.1 Overview of codified hashtags data set

Codified hashtags data set is composed of all tweets retrieved through the TwitterVigilance platform responding to the querying parameters listed in table 4.1. Monitoring period began on July 1st 2015 and ended on June 30th 2016. As shown in 6.2 a total of 25.185 tweets were collected, distributed in 7.569 native tweets and 17.616 retweets. A first characteristic of the data set is that 70% of messages are retweets, a very typical characteristic of Twitter use during critical situations. A high retweeting rate is recognized, in fact, as typical behavior on social media during a disaster event, responding

Table 6.2: Main features of Codified Hashtags channel

Total tweets	Native tweets	Retweets RT	RT % on total tweets	Unique Users	Active users (native tweets)	Active users % on total
25.185	7.569	17.616	70%	6.674	1.402	21%

Table 6.3: Codified Hashtags channel: temporal activity metrics

Total tweets	Days monitored	Days of activity	% of active days
25185	363	311	86%

to people's need to make information available [27, 79]. It is no surprise, considering that this channel is by definition composed of hashtags meant for warnings and emergencies to improve information sharing.

Unique users are 6.674 but only 21% of them are "active users", writing original tweets. Around 80%, instead, are participating only by sharing tweets.

As shown in figure 6.1 the great amount of messages is related to few high impacting events that occurred in summer and fall 2015. The highest peak of the channel during the period is related to Sardinia floods occurred on October 1st 2015, with 3189 tweets collected by TwitterVigilance Codified Hashtags channel. Peaks are characterized by orange color, which in TwitterVigilance is associated to retweets.

To measure how much the channel was active, we calculated an Activity Rate, considering as Active a day with at least one tweet published. As shown by table 6.3 there are 311 days out of 363 showing some activity; this means that the channel has an Activity Rate of 86%. Since tweets are connected to warning emissions, it is clear that a 100% rate is not achievable, even if messages may be posted also during "quiet period" to raise awareness and inform people for example about risks and life savings behaviors.

Other metrics are reported in table 6.4. It's worth to notice that compared to the total number of tweets and retweets there is a quite small number of **Replies**. This could be interpreted as a signal of a use of Twitter more as a broadcast medium than a conversational one. This is typical, for instance, of many institutional accounts which tend to use Twitter as a

Table 6.4: Main metrics of the Codified Hashtags channel

Favourites count	Replies count	Mentions count	Hyperlinks count	Hashtags count	Platform RT count *
9.768	164	22.748	18.052	58.314	329.478

channel where to publish press release instead of using it as an opportunity for real interactions with citizens [121, 123].

Hashtags more than double the number of total tweets. On average there are two hashtags within each tweet. If we look at the most used hashtags, the highest frequency is reported for some of the most used codified hashtags (part of the querying parameters of the channel). As table 6.5 shows, within the first twenty most used hashtags there are also names of geographic locations or words related to bad weather (like "maltempo" or "storm"), meaning that the codified hashtag is often used in combination with more simple tags indicating specific locations or particular weather events.

6.3.2 The Fab Six: the most used hashtags

One of the aim of this work is to verify if the proposed codified hashtags have been actually used, and furthermore, assess if in certain regional contexts data allow to prefigure a more stable adoption. By the analysis of the hashtag data set as represented in figures 6.2 and 6.8, it is very clear that few hashtags have been thoroughly adopted. In the majority of regional contexts they were poorly adopted or not adopted at all. Among the regions lacking in usage there are small ones, with limited geographic extension and scarce population, but also wide and important regions like Lazio, Lombardy, Campania or Apulia where the tweets collected are poor. Because the hashtag was used mainly in case a warning was issued, table 6.11 reports the number of warnings issued by the Central Functional Center of the Italian National Department of Civil Protection during the monitored period. The warnings are classified by color code, from yellow to red, and are related to rainfall impacts, namely hydrological risk (flooding) and landslide risk. Evidently, the number of severe warnings issued during the period is a first important indicator to evaluate the hashtags diffusion and the corresponding volume of tweets. Considering the red alerts emitted, as shown in table 6.11, Calabria had seven red alerts, Sardinia, Basilicata, and Sicily had three for each. Two

Table 6.5: List of the twenty most mentioned hashtags of the codified hashtags channel

Hashtag	**Frequency**
#allertameteotos	7870
#allertameteosar	6027
#allertameteocal	3580
#allertameteolig	3115
#maltempo	2553
#olbia	2039
#meteo	1784
#temporali	1528
#allertameteoer	1445
#sardegna	1091
#calabria	1055
#firenze	988
#allertameteosic	977
#rossano	916
#allertameteots	876
#protezionecivile	867
#idrogeologico	749
#toscana	696
#cfr	657
#prato	631

Figure 6.1: Time distribution of tweets in Codified Hashtags channel. Tweets are in blue, RT in orange

red alerts were issued on Abruzzi and one on Apulia, Liguria, Marche, and Molise.

For the purposes of this work, we considered the data sets of codified hashtags with the greater volumes of retrieved tweets. We define furthermore that a hashtag is adopted in a certain region if the number of unique users engaged in the hashtag-community is above the value of 278, that is the average value for the twenty codified hashtag channels.

Following this criterion, the most used codified hashtags were:

1. #allertameteoTOS, that is the top one with 7841 tweets collected in the monitored period and 1398 unique users;
2. #allertameteoSAR, the second one with 5977 tweets but it counts on the highest number of unique users 2297;
3. #allertameteoCAL, that is the third one with 3549 tweets and 1670 unique users;
4. #allertameteoLIG, with 3154 tweets collected in the monitored period and 743 users;
5. #allertameteoER, with 1447 tweets collected in the monitored period and 457 users;
6. #allertameteoSIC, with 980 tweets collected in the monitored period and 401 users.

Piedmont may be considered as a limit-case. The volume of tweets collected is half the amount of Sicily (406) but it is still higher than that reached in many other regions. But in Piedmont the number of unique users engaged is only 151. Compared to Sicily, Piedmont had far less warnings, only 2 orange warnings compared to 16 orange and 1 red warnings issued for Sicily during the considered period.

Each of the 6 most used codified hashtags was processed for main analytics. Results are interesting to assess different patterns of adoption of the hashtag in the regional contexts.

6.3.3 Pattern of usage

Looking at main metrics of the most used codified hashtags (see table 6.6) we notice that Tuscany, Sardinia and Calabria are the regions where the codified hashtag was more adopted. Some interesting differences in use emerge.

While Tuscany is the region where the codified hashtag (#TOS) reached the greater number of occurrences, Sardinia (#SAR) is the one with more **unique users**, considering both native tweets authors (Unique authors of original tweets) and authors of the whole stream (TW+RT). Sardinia and Calabria differ from Tuscany also for higher retweets occurrence (79% and 80% of total tweets) and a much lower tweets publication rate per authors, less than 3 tweets per author in Sardinia and Calabria, compared to more than 12 tweets per unique authors in Tuscany. Also the **Activity rate** varies greatly: with a 70% of active days in Tuscany compared to 24% and 29% in Calabria and Sardinia.

These elements identify different patterns of the use of codified hashtags in the analyzed contexts. For Sardinia and Calabria the majority of tweets were published during a few days when occurring severe weather events. The two regions were in fact troubled by devastating floods: the 1st of October 2015 in Olbia (Sardinia) and the 24th of August 2015 in Rossano (Calabria). The high number of tweets in Tuscany, on the contrary, is the product of a regular use of the hashtag during the whole monitored period, as proven by the activity rate at almost 70%. This pattern is also straightforward visible looking at figures 6.3 and 6.4, temporal distribution of tweets and retweets for Calabria and Sardinia, and figure 6.5 Tuscany. The data set of codified hashtags in Liguria (#LIG), Sicily (#SIC) and Emilia Romagna (#ER) present a suitable number of tweets, sign of a less diffused use. The ratio of tweets per user is around 6 in #ER, 4 in #SIC and 5 in #LIG. Activity rate is around 30% -35% for #LIG and #ER but only 11% in #SIC. See also table 6.8 for the whole set of hashtags.

6.3.4 Users in different contexts

Another point of the analysis was to gain insights about the type of accounts that mostly used the codified hashtags. On this purpose within the subset of the six most used codified hashtags (#TOS, #LIG, #SAR, #SIC, #LIG, #ER) we selected the first 100 mostly active unique users, those contributing with more original tweets. To better describe the channel and identify the communication pattern of different categories of users a manual annotation was made of this set of 100 top authors. Authors were classified into main categories, as described in 4.3.2. It was also considered the category "Bot" for accounts whose tweets were posted by an automatic publication software. These top 100 accounts were also manually annotated for **regional**

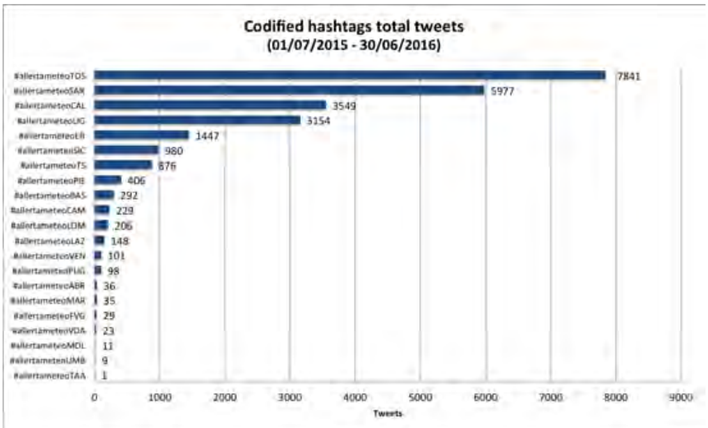


Figure 6.2: Bars representing the number of tweets published using each codified hashtag

Table 6.6: Main metrics for the six most used codified hashtags

Features per regional subset	**#SAR**	**#CAL**	**#TOS**	**#LIG**	**#ER**	**#SIC**
Unique Users (native tweets)	443	234	258	213	90	67
Unique users (TW+RT)	2297	1670	1398	743	457	401
Full Tweets	5977	3549	7841	3154	1447	980
Native TW	1250	709	3111	1036	510	299
RT	4727	2840	4730	2118	937	681
Native TW/users	3	3	12	5	6	4
Activity rate	29%	24%	69%	30%	35%	11%
RT/total	79%	80%	60%	67%	65%	69%

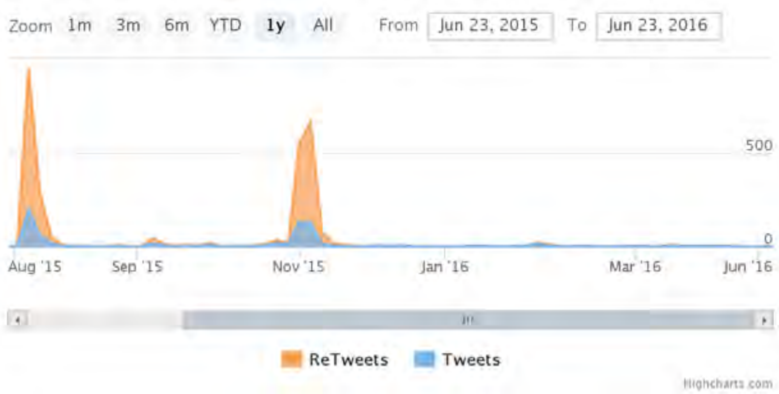


Figure 6.3: Temporal distribution of TW/RT for #allertameteoCAL

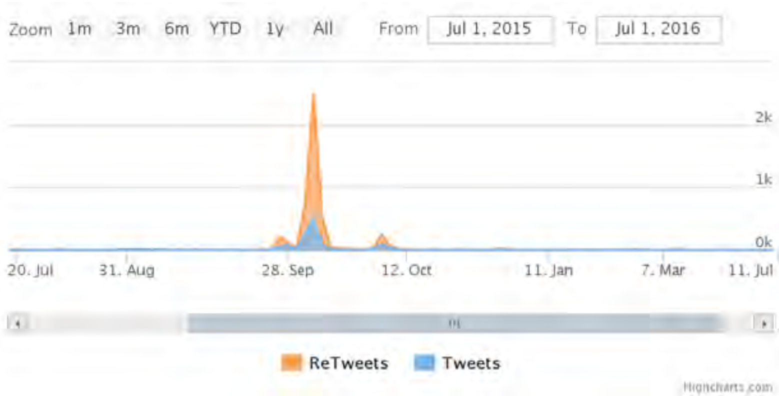


Figure 6.4: Temporal distribution of TW/RT for #allertameteoSAR

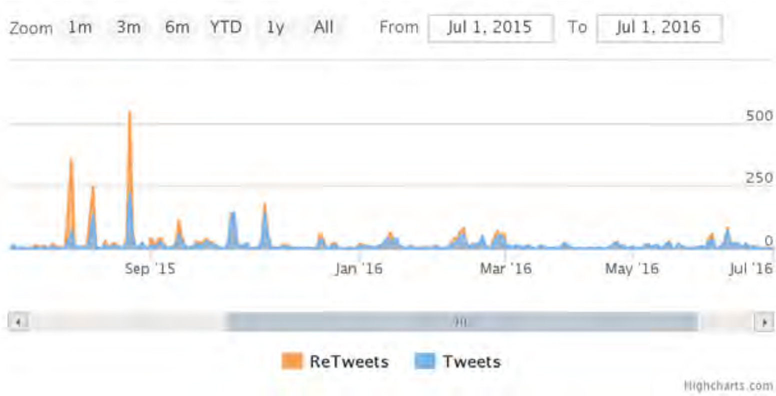


Figure 6.5: Temporal distribution of TW/RT for #allertameteoTOS

attribution. Almost 37% of top users are form Tuscany, another proof that the adoption of the codified hashtag has taken roots especially in Tuscany. Several users are also from Liguria and Sardinia. Accounts with a nation-wide dimension are 11%, and they are mainly Media accounts and a few commercial forecasting service accounts.

The annotation of tweet authors was also helpful to identify the most active categories of users: among the most active shows up Institutions, (27%), followed by Citizens (26%) and Media (18%). Looking more in depth at Institutions, it is possible to notice that half of the very active institutional accounts are from Tuscany.

This explains why codified hashtags in Tuscany have been adopted on a regular basis, due to the active role of institutions. Greatly supported by the account of the Tuscany public weather service (@flash_meteo), institutional accounts promoted the use of the codified hashtag to communicate and share weather warnings (see also section 1.1.5). Regularly adopted, the codified hashtag widespread to local institutions and volunteers in charge of civil protection and was integrated into official warnings communication. Institutional use pushed also the adoption of codified hashtags by media accounts and citizens. This pattern of adoption emerged only in Tuscany.

Mentions Similar considerations on users may be done looking at the most mentioned accounts, which were manually annotated following the same criterion. Accounts from Tuscany were the most mentioned, soon after

Table 6.7: Geo annotation of the 100 most active users (authors and mentions)

Regional annotation	N. of authors	N. of mentioned users
Tuscany	37	21
Liguria	12	14
National	11	24
Sardinia	11	14
Emilia Romagna	8	8
Calabria	4	8
undefined	4	3
Lombardy	3	2
Friuli V.G.	2	2
Piedmont	2	1
Sicily	2	2
Basilicata	1	0
Campania	1	0
Lazio	1	1
Puglia	1	0

national-interest accounts. These last ones are mainly accounts of media and institutional representatives, like the Italian Prime Minister @matteorenzi or @graziano_delrio, Minister of Infrasctructures, or @robertapinotti, Minister of Defense who were mentioned by citizens in tweets. The regional distribution of the most mentioned users is very similar to those of the most active unique authors, confirming the six regional areas where the codified hashtags were used the most. Among the most mentioned users, Institutions were even more relevant for mentions (accounting for 41% of top 100). The majority of these institutional accounts are from Tuscany that represents the most active context. Many are also the institutional accounts of Ministers, politicians or parliament members probably mentioned in a sort of call to action (or criticism) made by citizens.

Differences of six regional contexts Considering the top users of the six regional contexts where the codified hashtags were used the most it is possible to have further elements to better understand hashtags adoption.

The bars of fig 6.6 indicate that in Tuscany, Emilia Romagna and Liguria

Table 6.8: Users' category of the top 100 accounts (as authors and mentioned)

Users category	**N. of unique authors**	**N. of mentioned users**
Institutions	27	41
Citizens	26	22
Media	18	21
NGO	14	5
Weather	12	11
BOT	1	-
Not identified	2	-

many accounts of Institutions show up among the first 20 most active users, less in Sardinia and Calabria, none in Sicily. In Tuscany the numbers of tweets published by the most active institutional accounts is far higher compared to other regions; Media are quite active in all contexts; Citizens are more active in Calabria and Sardinia as an outcome of participation during high impact events happened there in 2015. Even if #SAR and #CAL are the most numerous data sets after Tuscany, as shown in table 6.6, the top 20 users in #LIG published more tweets, especially Citizens: a sign that in Liguria, as an outcome of previous flooding events (see 5) citizens became very active in using Twitter, also during non-emergency periods.

Also the analysis of the most retweeted authors in each data set confirms this findings, as shown in tables published in Appendix A. Data show how in Tuscany, but also in Emilia Romagna and Liguria, there is an important occurrence of institutional accounts among the ten most retweeted users of the channel, while in Sardinia the ten most retweeted users are all Citizens.

6.4 Diffusion of the codified hashtags

We also tried to find out how many people used the codified hashtags compared to potential users. The best way to give an answer has been to compare the volume of codified hashtags to the volume of other tags emerging during high impact events (like #alluvioneRossano or #alluvioneOlbia). When the monitoring channels in the TwitterVigilance platform were set up, we

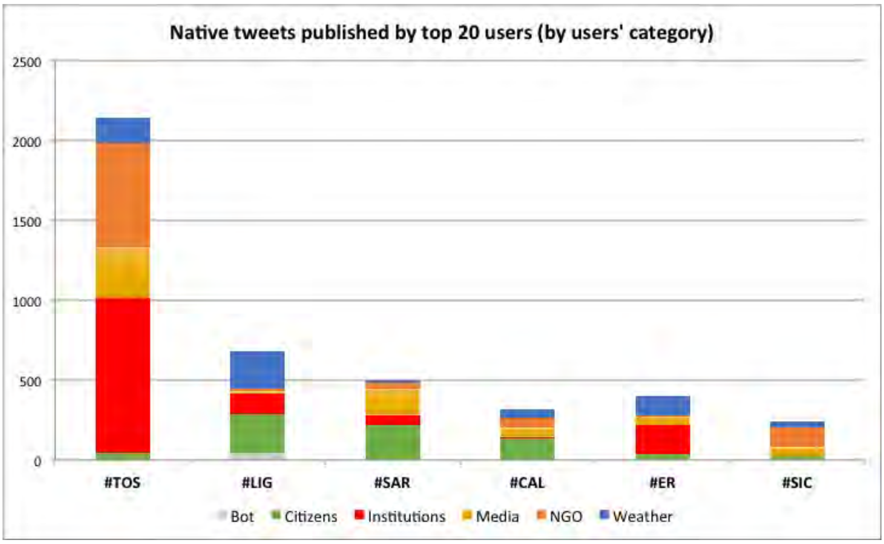


Figure 6.6: Native tweets by the top 20 users of each regional hashtag (by category of users)

Table 6.9: Main features of the monitored channels

Channel	**#bad weather**	**#weather**	**Forecasters Users**	**Codified Hashtags**
Total tweets	1.843.095	740.758	124.148	25.185
native tweets	1.228.332	519.830	45.308	7.569
retweets	614.763	220.928	78.840	17.616
TW % on total	67%	70%	36%	30%
Unique users (TW)	222.813	56.507	5.524	1.402
Unique users (Tw+RT)	335.831	116.370	18.328	6.674
Active users %	66%	49%	30%	21%

decided to create three different channels to be used as complementary domain for the analysis. These channels are aimed at retrieving tweets about weather, in general, and bad weather conditions, in particular. A description of these channels querying parameters is provided in section 4.2.2.

A comparison between the codified hashtags data set and these reference data sets was performed.

Table 6.9 presents the main features of the data sets related to messages retrieved during the monitoring period: form July 1st 2015 to June 30th 2016. This table gives us an idea of the number of people that has been using Twitter to share messages about forecasts, following main forecasting services or simply commenting the weather. The bigger one is channel #badweather (BW), followed by #weather (WW) and @Forecasters users (FU). All these data sets are much bigger compared to the codified hashtag channel. With nearly 2 million tweets, BW channel is 73 times bigger than the codified hashtag channel, WW is 29 times bigger and FU, the smallest of the three, is anyway 5 times larger. On average this means that around 5 thousands tweets were sent daily about bad weather conditions. A first interesting thing to notice about these channels is that BW and WW are mainly expressive channels having a 70% of tweets and only a 30% of retweets. A distribution that is the opposite of CH and FU channels. A high retweeting rate is typical of informative channels where users tweet with the aim of sharing useful information. Furthermore, looking at the ration between Active Users, the percentage of users writing original tweets, and the whole amount of users involved in the channel, BW and WW have a higher rate than FU and CH. This confirms that the former are expressive channels, collecting tweets of people talking about the weather as happens on the bus or inside the elevator, while the latter are informative channels where users employ the hashtag to share news, reports, advice and facts about the weather. During an emergency could be very difficult to find useful information in those kind of expressive channels. Loads of this kind of tweets may not be of any use and might not increase situational awareness during a critical situation. High volumes in tweets in this generic channel may not be in relation to critical events. For example looking at fig 6.7 we may see that the highest peak in daily volume of tweets is on August 10th 2015, when several thunderstorms hit many cities in the North and the Central part of Italy. Even if some cities reported local damages due to storm effects, it was generally a less severe situation compared to Sardinia or Calabria floods, which are the highest peaks in daily volume of tweets for the CH channel. Probably the high volume of tweets on August 10th is related to the fact that bad weather conditions were spoiling the central week of the Italian summer holidays, with many users complaining for rain showers at the seaside. The highest peaks of daily volume of tweets in CH channel are both Sardinia (3189) and Calabria (1163) floods. October 1st is also the highest peak in WW and FU channels, respectively reaching 6354 and 736 in daily volume of tweets.

6.4.1 Spreading of codified hashtags in weather related channels

Among these expressive channels we calculated how many times the codified hashtags were mentioned in any of the other three channels.

Table 6.10 reports for each channel how many times each codified hashtags was mentioned (table shows only the most widespread hashtags). Generally, we may say that codified hashtags hardly spread into other weather related channels collected through the TwitterVigilance platform during the considered monitoring period. Hashtags which spread the most were #TOS, #SAR, #LIG and #CAL. The codified tags were more mentioned within the WW channel, but one of the reasons could be the similarities in querying parameters. In fact, in all codified hashtags is included the word "meteo" that is the querying parameter for WW channel, that in many cases is also part of the Twitter accounts of many forecasting services (like @3Bmeteo or @flash_meteo).

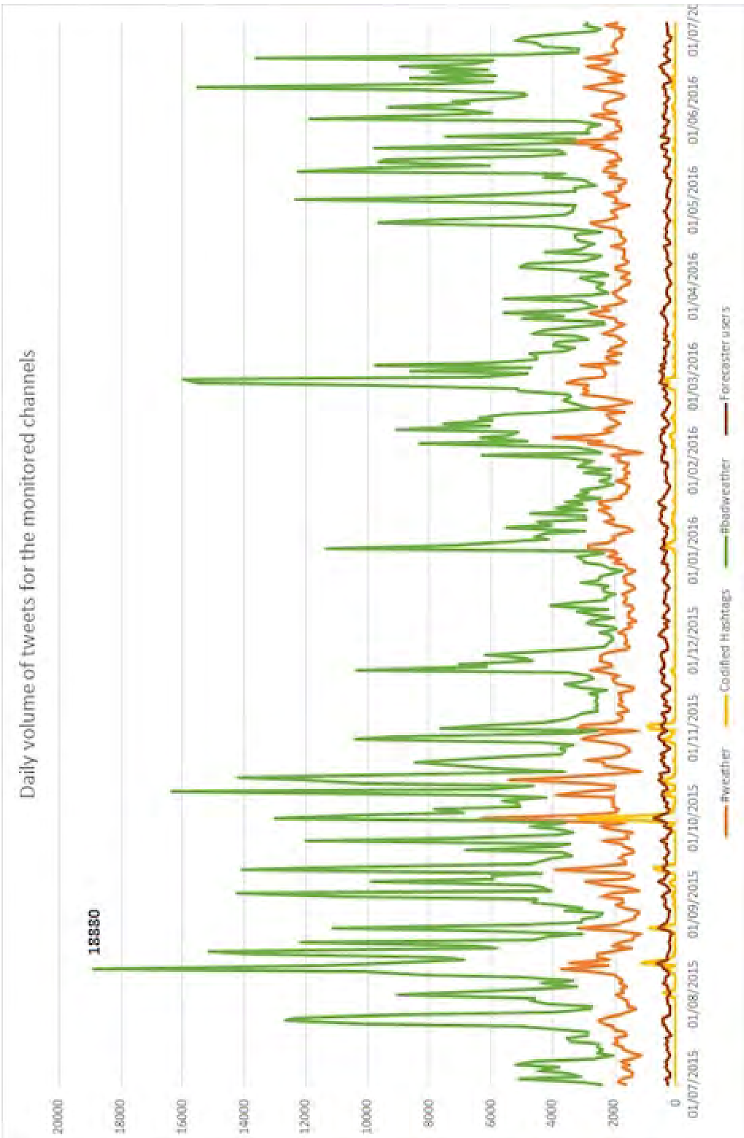


Figure 6.7: Daily volume of tweets of the channels

Table 6.10: Spreading of codified hashtags in weather related channels

codified hashtags	**BW channel**	**WW channel**	**FU channel**	**CH channel**
#TOS	3549	7781	1548	7841
#SAR	1272	5286	240	5977
#LIG	1047	3139	95	3154
#CAL	660	2942	80	3549
#ER	547	1400	179	1447
#SIC	537	977	65	980
#PIE	181	374	11	406
#LOM	128	194	27	206
#BAS	123	286	0	292
#LAZ	86	155	49	148
Total sum	8442	23976	2307	25462

6.5 Conclusions

Analyzing the data collected with the help of the TwitterVigilance platform during the monitored period (from July 2015 to June 2016) it emerges that only six out of the twenty proposed codified hashtags for weather warning showed a significant adoption. Tweets including the Tuscany codified hashtag #allertameteoTOS account for more than the 30% of the whole Codified Hashtags data set. A comparison of the different contexts allowed to highlight the differences in pattern of usage. Tuscany appears to be a case of "**regular use**", as confirmed by the highest percentage of activity rate (69%), the highest number of collected tweets and the highest ratio of tweets per user (12). No red alert was issued for Tuscany during the monitored period, while the orange warnings were 13, far less than many others cases, like for instance Campania, Lombardy, and Apulia. In Tuscany the codified hashtag was actually adopted to communicate every warnings (also the 63 yellow warnings issued) and institutions had a primary role in building the hashtag-community.

At the opposite, Sardinia and Calabria show a kind of "**burst use**" related to exceptional and isolated situations linked to the occurrence of a disaster. They count on the highest numbers of engaged users and high number of tweets but only during the limited days of the emergency. Activity rate is in fact around 25-30% during the whole monitored period. Twitter activity around the hashtag is pushed by extraordinary circumstances and is largely sustained by Citizens. These are also the contexts where Institutions appear less engaged in hashtag adoption and in Twitter in general.

Emilia Romagna and Liguria show an "**tentative use**", activity rate is about 35% indicating that the community is not yet fully engaged; temporal distribution of messages is more regular compared to Sardinia and Calabria. Institutions are not fully engaged in this adoption but they have started. In Sicily adoption appears weaker, with only 11% of activity rate. Anyway, this is a very limited activity considering that Sicily was interested by the emission of 3 red and 13 orange weather warnings. For Piedmont the volume of collected tweets was smaller than in the previous contexts, but the number of warnings was also limited. Because of the small amount of published tweets, Piedmont was not considered as a case of consistent adoption in this work. But we may assess that this is a context where the attention for the use of the codified hashtag is increasing, also due to the support of a few institutional accounts, like the Piedmont regional government or the Civil Protection of Alessandria (@PCProvaAL). This last one organized special training for its personnel, volunteers and for those citizens interested in becoming digital reporters during the emergencies. The attention to codified hashtag use is also confirmed by what happened during the flooding events of November 2016 (not covered in this analysis) where the codified hashtag #allertameteoPIE was used in 7000 tweets in a few days.

In other contexts hashtags appear poorly used during the monitoring period. Contexts like Basilicata and Lazio seem at an "**embryonic stage**", with a tentative use proposed by some users, but not really regularly embraced. Other regions, interested by several orange warnings during the considered period did not show any use of hashtags, like Lombardy, Campania or Apulia. Of course many different reasons can explain this diversity, starting from population differences, geographical digital divide, differences in social media use in urban and rural contexts. One reason could also be the different regional climatic conditions. Regions on the Tyrrhenian Sea, like Liguria, Tuscany, Calabria, Sardinia, Sicily are more exposed to exceptional rains and consequent flash floods or major flooding. In these contexts, institutions and citizens probably faced more often this kind of emergencies and turned to social media as a new way to cope with this events. Of course, an important reason is related to the different use of Twitter by key players, like the regional weather services or the local institutions responsible for the warnings. As reported in chapter 3, weather alerting in Italy is organized on a regional basis through the Decentralized Functional Centers. The National Department of Civil Protection has no Twitter account, and not all the regional centers have one. There is a specific account for weather update only in few regions (see table 3.2, around 6 at the time of this analysis). Only few of them use the codified hashtag. The greater uptake of some regions respect to others during the monitored period may be explained by two elements: the uptake is achieved only when institutions are active on Twitter, engage local users and create certain communication habits, also with formal documents, like it happened in Tuscany. Alternatively, the uptake emerged only during disaster events, like Calabria and Sardinia, where the codified hashtag is sustained by citizens (see chapter 7 for case studies). In this case, though, a true uptake is granted only if influencers or media accounts get involved sharing the hashtag thorough their vast networks of followers (in Sardinia the most retweeted author has more than 110.000 followers, as reported in Appendix A, table A.2).

Digital communication rules for emergency situations need to be shared with the population and should represent the different contexts. The codified hashtag may not be the only solution, but the solution needs to be participated by the society at the local level.

Table 6.11: Number of weather alerts issued during the monitored period in the different regions. It reports the number of warnings issued by the National Functional Center of Civil Protection in Rome. The warnings are only related to hydrological risks like floods and landslides.

	Yellow	**Orange**	**Red**
Calabria	139	33	7
Abruzzo	136	12	2
Molise	112	9	1
Lombardy	106	24	0
Umbria	104	4	0
Apulia	98	21	1
Marche	94	6	1
Lazio	87	10	0
Basilicata	81	18	3
Sicily	79	16	3
Campania	75	29	0
Veneteo	70	9	0
Emilia Romagna	66	7	0
Tuscany	63	13	0
Liguria	46	4	1
Piedmont	46	2	0
Friuli Venezia G.	37	2	0
Sardinia	26	4	3
Trentino A.A.	17	5	0
Valle d'Aosta	7	0	0

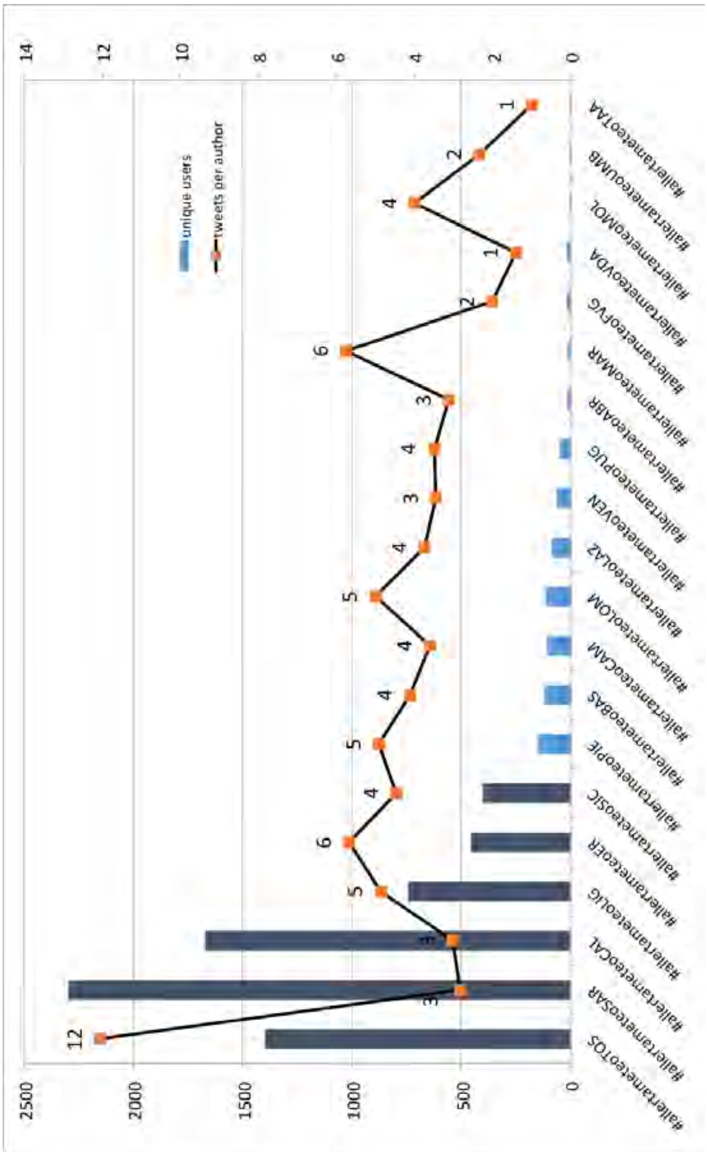


Figure 6.8: For each codified hashtag are presented the number of unique users and the average amount of original native tweets per user.

Chapter 7

Case studies

In this chapter we present an analysis of case studies of severe weather events to assess if and how codified hashtags have been used in critical situations. Events were selected as to cover different regional areas and allow a comparison of diffusion level in those contexts. The first two events are related to disastrous flooding in the South of Italy, the third one is a less impacting event involving great part of the Italian peninsula and in particular Central and Northern regions. This last case helped to understand whether the few occurrences of certain codified hashtags can be attributed to a lack of impacting events or rather are the result of a poor employment of the hashtag. For the selected events a Social Network Analysis was also used to model the propagation of information and identify influential users. The analysis also evaluated a few key Twitter metrics of the more generalist channels monitored via the TwitterVigilance platform.

7.1 Event selection and methodology

To make further evaluations on the codified hashtag use we investigated its adoption during selected events. Three high-impact events were chosen by looking at peak volumes of daily tweets collected in the Codified Hashtags channel. The highest volumes of daily tweets in the data set were recorded on Oct. 1st 2015, at top level with 3189 tweets (554 native tweets - 2635 RT); Aug. 12th with 1163 messages (198 original tweets and 965 RT); September

30th with 1016 messages (257 original tweets and 759 RT). The first and the third belong to the same event, the Sardinia floods, the second one is related to the Calabria flood. A third event was selected by choosing among severe weather events covering different regions of the North of Italy, and this intending to verify if the low usage of codified hashtags resulting from major analysis, (see Chapter 6), was a result of a lack of adoption or of impacting events. We summarize here the main features of the selected events:

1. Sardinia was hit in fall 2015 by a violent storm that caused several floods. Analyzing the event, we selected three days ranging from Spt. 30 to Oct. 2 2015. During these days the 76% of all #allertameteoSAR tweets collected over the monitored period was published.

2. In Calabria a violent downpour caused floods in several parts of the Ionian Calabria, in particular the city of Rossano was dramatically flooded. The selected event covered a day range from Aug. 10 to 12. In these days the 42% of whole tweets volume of #allertameteoCAL was published.

3. Bad weather struck Italy in different areas of the Center and the North at the end of February 2016. Widespread rainfalls and downpours covered Liguria, southern Lombardy, Veneto, Friuli Venezia Giulia, Emilia Romagna, Tuscany, Sardinia, Umbria, Lazio, Abruzzo and Molise. Selected monitoring days range from Febr. 28 to Mar. 1, 2016.

For each case study we analyzed main Twitter metrics following methodology proposed in Chapter 4. To model information propagation in the specific events we used Social Network Analysis approach as described in section 4.3.3.

7.2 Case study 1: Sardinia floods

7.2.1 Meteorological scenario

Sardinia was hit on Oct. 1st by a violent storm that caused torrential rains over great part of the region. The low-pressure system widely struck the Cagliari area where severe flooding occurred, then moved northward hitting hard the Gallura area and Nuoro. Also in Olbia torrential rains caused severe flooding. The Mayor advised citizens not to leave their houses. Due to the very heavy rainfalls, canals overflowed in rapid succession: first Rio Siligheddu, then Rio San Nicola and Rio Gadduresu; all of them already overflowed two years ago. Serious problems were reported also in Nuoro, in particular in the TorpÃ¨ area where the incessant rain blasted a cemented canal running through the village,inundating it. Forty families were displaced, while another group, of four, remained isolated due to the collapse of a bridge. The Central Functional Center of the Civil Protection issued a red warning for heavy rains and floods for Sardinia; orange warning for flood risk was issued for Sicily and yellow warning for heavy rain was issued for Central and Southern regions.

7.2.2 Tweets analysis

Among the Tweets collected by the TwitterVigilance platform in the codified hashtags channel, we analyzed those published during the period from Sept. 30 to Oct. 2. Tweets collected on the CH channel in these three days amount to 4972, with 946 native tweets and 4026 retweets. Retweets are 80% of the whole set of collected messages, confirming the exceptional nature of the event [29, 177]. Retweets count as provided by Twitter platform amounts to 120.840. Tweets were published by 2006 unique users, 80% of them contributing only by retweeting. Tweets and users reached their highest peak on October 1st, with 66% of collected tweets on that single day.

Looking at the ten most used **hashtags**, (see table 7.1) many are names of geo-locations and places involved in the flood, like #Olbia, that is the second most used, and #Sardegna, the third one, or #alluvioneOlbia (Olbia flood) composed with the typical syntax for local disaster: type of event and place. We find also other hashtags created with a similar syntax like #emergenzaolbia (26 occurrences) and #maltempoolbia (25 occurrences). Within the ten most used hashtags also show up simple words semantically related to the event, like *temporale* (storm), *maltempo* (bad weather conditions) and *meteo* (weather). Hashtags analysis is important to identify trending terms. It also provides important information that may used for recommendations mechanism system automatically promoting the codified hashtag towards other groups of users.

Figure 7.1 shows very well how during the event we can recognize a synchronization of the relative maximum of tweets containing the different hashtags and semantically related words. The maximum volume of tweets, the highest curve, is for the codified hashtags followed by the name of the

Figure 7.1: Synchronization of semantically different hashtags during the Olbia flood, 1st October 2015.

Table 7.1: Top 10 hashtags used during the Sardinia flood - CH channel

Hashtags	**Frequency**
#allertameteosar	4584
#olbia	1863
#sardegna	639
#maltempo	385
#meteo	272
#temporali	235
#allertameteolig	219
#cagliari	193
#allertameteotos	93
#alluvioneolbia	92

place impacted by the event, Olbia in this case. This pattern is also recognizable in other events see also figure 7.4. Within the top ten hashtags appear also codified hashtags of other two regions, #allertameteoLIG and #allertameteoTOS, where yellow warnings were issued during the selected period, and users where using the codified hashtag in tweets.

Total **hyperlinks** retrieved in the data set of codified hashtags for the monitored days are 10207, corresponding to 394 unique URLs. Considering only the 946 original tweets published during the three days, URLs are 394.

Favorite tweets account to 2992, 63% (1903) of which recorded on Oct. 1.

Concerning the category of **users**, Citizens resulted to be the more engaged around the Sardinia codified hashtag during the event. The widepsread dissemination of hashtags was in fact mainly done by influencers like @insopportabile (with more than 100.000 followers), who actively participated sending more than 60 tweets during the peak days (see table 7.2). He is also the most retweeted unique user. Institutional accounts appear only if we look at authors of native tweets, where among the top 20 authors we find Comune di Cagliari, Comune di Oristano and the Civil Protection of Florence.

The codified hashtag #allertameteoSAR spread out and was very used; even the Minister of Defense used it in tweets. The most **favorite** message was the following one by @robertapinotti, the Italian Ministry of Defense (104 RT and 90 likes)

Table 7.2: #allertameteoSAR: top 20 authors and coded category. Full and native tweets.

Top 20 authors (TW + RT)	category	freq	**Top 20 authors (TW)**	category	freq
ecatetriformis	citizen	95	sardegnaoggi	media	31
insopportabile	citizen	65	emergenza24	NGO	29
paoloigna1	citizen	63	giamplus	citizen	22
alienakent	citizen	52	insopportabile	citizen	21
llimantul	citizen	50	alienakent	citizen	20
emergenza24	NGO	47	egyzia	citizen	20
criskikka75	citizen	40	youreporter	Media	19
elisabetta_sann	citizen	40	robyzoc	citizen	18
eddiepepsi	citizen	38	emergenza24pro	NGO	17
gav_gavi	citizen	38	paolomastino	citizen	16
gabriele81ferra	citizen	37	nicola_pinna	citizen	15
margotten78	citizen	32	wwwmeteoit	weather	14
sardegnaoggi	media	31	abbanoa	citizen	12
giamplus	citizen	30	franci_orgiana	citizen	12
torillo59	citizen	28	guidopiga	citizen	12
egyzia	citizen	26	comune_cagliari	institutions	10
robertore62	citizen	26	lafrogh	citizen	10
robyzoc	citizen	24	medherald	citizen	9
cambagigi	citizen	23	protcivcomunefi	institutions	9
18undici	citizen	21	comuneoristano	institutions	8

> "*Stiamo inviando uomini e donne del 5 RGT genio di Macomer per fornire aiuto alla popolazione di Olbia colpita dal maltempo #allertameteoSAR*". (We are sending men and women of the 5th RGT Macomer genius to deliver aid to the people of Olbia hit by bad weather #allertameteoSAR)

7.2.3 Network analysis

The engagement of different categories of users during the event was represented also by using Social Network Analysis. The figures 7.2 and 7.3 present two retweet graphs made with Gephi visualizing two different centrality measures. Figure 7.3 represents a retweet graph of communication propagation among users (a node is a Twitter user) during the Sardinia flood based on **eigenvector centrality** values. Node size is proportional to the computed value of eigenvector centrality (EC) of the node. Eigenvector centrality measures how much a node is influential in the network. This influence is calculated by a specific formula that gives relative scores to nodes assuming that nodes connected to high-scoring nodes get a higher score. As to say that a node is more influential if it is connected to influential nodes. Edges represent the retweeting pattern between two nodes. Figure 7.2 represents another centrality attribution that is **betweenness Centrality** (BC). Betweenness measures how often a node acts as a bridge along the shortest path between two other nodes in the network, connecting users groups that would be otherwise disconnected. The two centrality measures highlight different functions that users may have in propagation of information in the network: nodes with high EV value are the influencers, nodes with high BC are those connecting groups and thus making the information more viral in different communities. These last ones are very important for hashtags diffusion, helping to spreads its use in communities that are not at first directly aware of the hashtagging code for emergency.

To produce these graphs we used the Gephi layout *ForceAtlas 2*, with default settings. Given the high number of users, images present networks of unique users filtered by Betweenness centrality values, ranging from 1010 to 52102. For the other graph in figure 7.3 we filtered for eigenvector value ranging from 0.3 to 1.

Both graphs show that Citizens are the most influential actors in the network, all bigger nodes are in fact green, coding color for the Citizens category. In figure 7.2 the graph shows clearly that users acting as information hubs are Citizens, they are even more important than Media users (yellow). Among institutions important users are Cagliari Municipality and Sassari Municipality. Therefore, the codified hashtag in Sardinia was mainly diffused thanks to citizens (users name is labeled on the node) which acted as hub and influencers for propagation of information. The successful propagation of a hashtag is very much linked to the role played by authority or by influencers [58, 181]. Therefore, it is quite predictable that when hashtags are introduced by an influencer they will go viral. What we can notice is that this role was left to citizens as the institutions played a limited role on social media during the emergency. Only if institutions start playing a more active role it is possible to overcome possible problems arising from lack of coordination among online volunteers and official agencies. Social media may help to increase the communication and interconnections between institutions and emergent groups of citizens committed to helping but when the coordination role is left to citizens this may raise an issue.

The graph on influential users, in figure 7.3, is very similar to the previous one: green nodes are the bigger ones, confirming citizens as main actors. We notice that even if the same users show up in the two graphs, bigger nodes are different, with citizens playing different roles in the network and everyone contributing to the hashtag propagation.

7.2.4 Codified hashtags diffusion

To assess hashtags diffusion we considered volumes of tweets published in selected days and collected in channels BW, WW and FU, with semantically broader filtering parameters (see section 4.2.2 for more details). Main features of the four channels were reported in table 7.3. BW channel (where retrieval was done by querying Twitter API for words and hashtags generally used to describe bad weather conditions) counts 13.020 tweets on the 1st of October, the day of the Olbia flood, that is four time bigger than highest volume collected in Codified Hashtags. WW channel (retrieving tweets containing hashtag #weather) amounts to 6.354 tweets, double of Codified Hashtag. Percentage of retweets in the different channels is though quite different: CH has 83% of retweets while BW only 38% and WW 52%. The low retweeting rate is typical of those more *annunciative* channels used by people to comment and share personal thoughts or emotions about the event and less used to deliver useful information [27, 29].

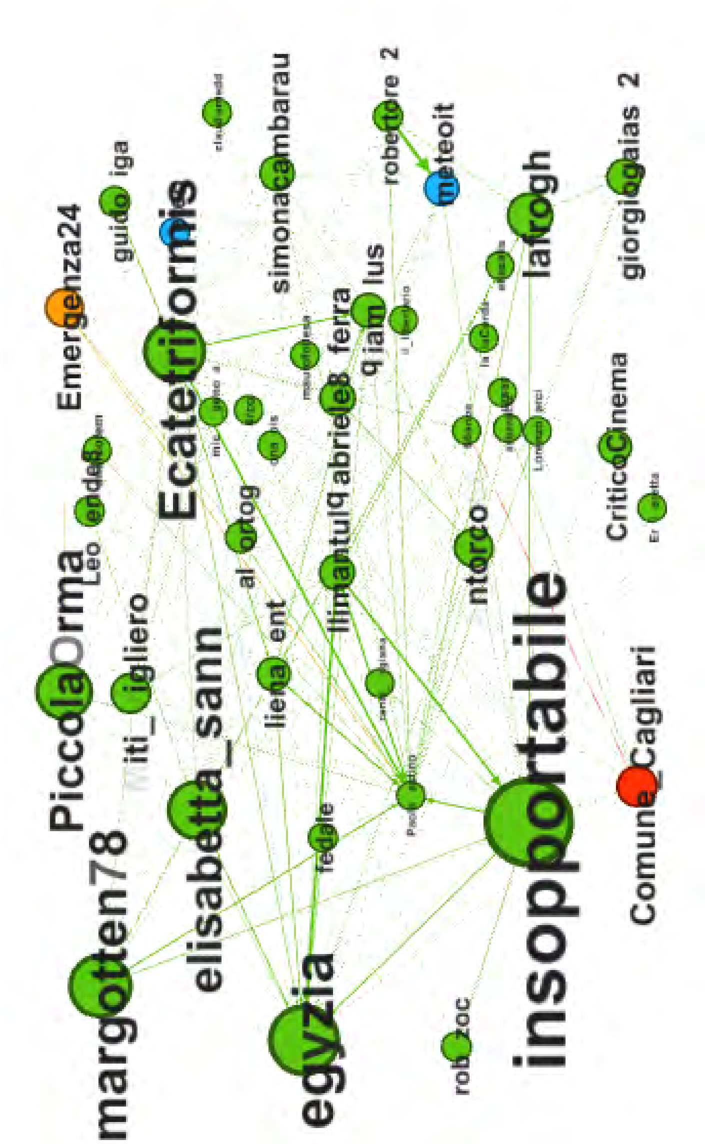


Figure 7.2: Retweet graph of codified hashtags during the Sardinia flood. Node size is proportional to the value of betweenness centrality. Node color relates to users category: red for Institutions; green for Citizens; orange for NGO; yellow for Media; blue for Weather services.

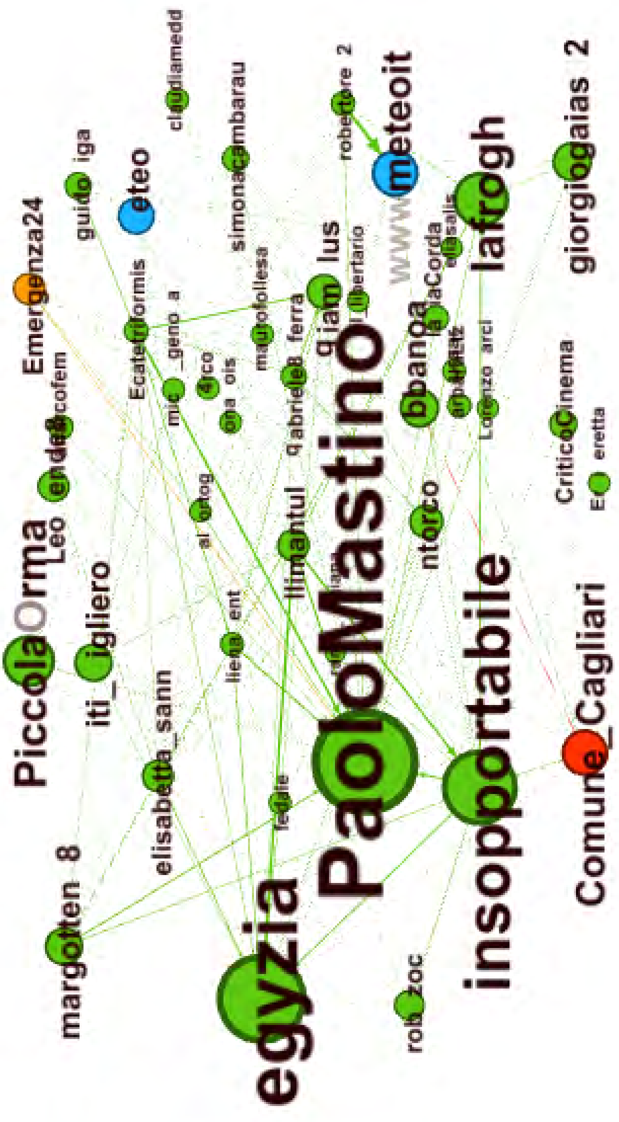


Figure 7.3: Retweet graph of codified hashtags during the Sardinia flood. Node size is proportional to the value of eigenvector centrality. Node color relates to users category: red for Institutions; green for Citizens; orange for NGO; yellow for Media; blue for Weather services.

Table 7.3: Main features of the monitored channels on October 1st, during the Sardinia flood

	RT + TW	TW	RT	RT % on total	Unique users
BW channel	13020	8030	4990	38%	8219
WW channel	6354	3060	3294	52%	2959
FU channel	734	195	539	73%	352
CH channel	3304	557	2747	83%	1453

7.3 Case 2: Calabria floods

7.3.1 Meteorological scenario during the Calabria floods

From August 9 to August 12 a severe meteorological configuration struck many parts of Italy. In particular severe events occurred in Tuscany and in the South of Italy, where a violent downpour caused a flash flood in Rossano, Calabria. In Tuscany a waterspout hit Marina di Carrara and a woman was seriously hurt. The upheaval also affected the coast of Livorno and Pisa: 6 persons were injured due by fallen trees.

On the night between the 11th and the 12th heavy showers and thunderstorms hit the Northern Ionian Calabria. Heavy rains relentless fell in the night near the Calabrian ionic coast, especially between Rossano and Corigliano Calabro with amounts that locally reaching 150-160 mm (approximately equivalent to 5 times the average amount of the entire summer) and caused flooding of houses and streets. A storm caused a flash flood in Rossano Calabro: hundreds of cars were dragged from streams full of mud and debris. In Rossano more than 10.000 people remained without electricity; five hundred people were displaced in Rossano, including locals and tourists.

7.3.2 Twitter analysis

The days identified for the retrieval of tweets through the TwitterVigilance platform on the codified hashtags channel were 9 - 12 August 2015. Tweets collected in the CH channel in the three days amount to 1827, with 417 native tweets and 1410 retweets.

Figure 7.4 shows the typical pattern with synchronization of relative peaks of tweets volume of the different hashtags. The geo-location term

Table 7.4: Top 10 hashtags used within CH channel during the selected event.

Hashtags	**frequency**
#allertameteocal	1.172
#rossano	771
#allertameteotos	563
#calabria	293
#alluvione	157
#corigliano	132
#meteo	131
#sudnelcuore	123
#toscana	95
#cosenza	89

Rossano reached the highest volumes of tweets.

Retweets are 77% of the whole set of collected messages confirming the exceptional nature of the event. Tweets were published by 947 unique users, 80% of them contributing only by retweeting. Tweets and users were highest during the peak day, Aug. 12, accounting for 65% of collected tweets. During the peak day retweets were the 83% of the whole set. Looking at the ten most used **hashtags**, (see table 7.4) soon after #allertameteoCAL we find #Rossano and #Corigliano, the village most impacted by the floods and #Calabria. The third most used is the codified hashtag for Tuscany that in the same days was hit by stormy weather on the coast where a whirlwind caused a lot of damage in Carrara. Among the most used hashtags also we find #allertameteoER, #allertameteoBAS and #allertameteoPUG (respectively mentioned 31, 28 and 21 times). Total **hyperlinks** retrieved in the data set of codified hashtags for the monitored days are 1362, corresponding to 346 unique URLs. Considering only the 417 original tweets published during the three days, URLs are 328. **Favorite** tweets are 891, 64% of which recorded on Aug. 12.

As for the category of **Users**, (see table 7.5) among the top 20 unique authors of native tweets (those contributing with original messages published during the monitored days), many of the Institutions and Media accounts appearing are mainly related to Tuscany and not to Calabria, scene of the flood. Also the account of the Calabria regional weather service @Cfm_arpacal ap-

Figure 7.4: Synchronization of semantically different hashtags during the Rossano flood, August 12th 2015.

Figure 7.5: One of the most retweeted messages published on August 12th 2015 within the Codified Hashtags channel

pears after Tuscany @flash_meteo. Among the most retweeted authors is the Calabria regional weather center, @Cfm_arpacal, second only to the account @Ossmeteobargone, a meteorological observatory in Emilia Romagna that is a very active account. @Cfm_Arpacal had however a central role in information propagation, as it is well showed in the retweet graph presented in the next section, in figure 7.6. Many citizens also appear among the most retweeted authors; people from Calabria shared information and photos during the flood. One of the most retweeted messages was published by @dalecooper30, on Aug. 12 ,who posted a picture of the flooded streets of Rossano; the message was retweeted 81 times and marked as favorite 32 times (figure 7.5).

It's worth to notice that no account of local institutions is mentioned

Table 7.5: Top authors by category on CH channels during the Calabria flood

Top 20 authors (native TW)	category	TW	Most retweeted authors	category	RT count
qn_lanazione	media	26	Ossmeteobargone	weather	712
emergenza24	NGO	22	Cfm_Arpacal	weather	456
cesipcrosignano	institutions	10	Emergenza24	NGO	383
controradio	media	9	geoluc_mei	citizens	314
intoscana	media	9	santosalfi	citizens	312
muoversintoscan	institutions	9	0Zenzy	citizens	285
damgiordano	citizens	8	FabioFtrifi	citizens	268
flash_meteo	weather	8	amtoscana	bot	260
valeria_sang	citizens	8	Emergenza24Pro	NGO	235
prov_fi_pc	institutions	7	mafaldina61	citizens	231
vabcollinemedic	NGO	7	TaniuzzaCalabra	citizens	197
iltirreno	media	6	GraziaTardo	citizens	195
michiamomita	citizens	6	sulsitodisimone	bot	171
protcivsestofi	institutions	6	MonacoCarla	citizens	159
taniuzzacalabra	citizens	6	fbg88	citizens	151
alfonsobombini	citizens	5	manginobrioches	citizens	151
Cfm_arpacal	weather	5	PDibattista	citizens	151
costantinigenny	citizens	5	vespro4	citizens	148
emergenza24pro	NGO	5	armandoscalia	citizens	143
lucalombroso	weather	5	Nazzi_Mi	citizens	137

among the top authors of table 7.5. This is also confirmed by looking at retweet graphs presented in figure 7.6 and 7.7.

7.3.3 Network analysis

To identify important users (nodes) in the network, like influential users or opinion leaders, we used Social Networks Analysis on the data set (see also section 2.3.2). Each category of users was coded with a color: red for Institutions; green for Citizens; orange for emergency NGO; yellow for Media; blue for Weather services. Figures 7.6 and 7.7 represent retweet graphs of communication propagation, regarding the codified hashtag during Calabria floods. Node sizes express the value of Eigenvector or Betweennes centrality. The graph was created using the Gephi software; the network layout employed is *ForceAtlas 2*, with default settings. Image 7.7 presents networks of unique users filtered by Eigenvector centrality values, ranging from 0.3 to 1; network in figure 7.7 is filtered by Betweenness values of minimum 1090; nodes with inferior value are not represented.

In figure 7.6 we see how bigger nodes, with higher Betweenness values, belong to different categories: Weather services, Citizens and NGO; users of different categories acted as important information hubs in the hashtag community. Two institutional accounts are visible but they are not local: one, @italiaSicura, is the username of a National Agency for the Prevention of Hydrological risk and the other one, @muoversintoscan, is a Tuscany agency for traffic and transports information. Also @cfm_Arpacal, regional Functional center for Calabria has a central role, well connected to NGO and Citizens. Considering the most influential users, with high value of Eigenvector centrality, figure 7.7 shows a different picture: @CFM_Arpacal is smaller than Tuscany weather service, @Flash_meteo; also many institutional accounts, red nodes, emerge, and it's interesting that the majority of them are related to Tuscany. The most influential users, biggest nodes, are NGO's accounts @anpasnazionale (Emergency and rescue NGO), and @Emergenza24, an association dealing with risk and emergency communication very active on Social Media at national level.

7.3.4 Codified hashtags diffusion

To assess how much the codified hashtags were used during the peak days, we made a comparison between volumes of tweets retrieved by the Twitter-Vigilance platform and more generals channels related to bad weather conditions (BW), weather (WW) and the channel retrieving tweets published by weather forecasting services users (FU). A first thing to notice is that the highest volume of tweets in BW channels is not reached on August 12th, the day of the Rossano flood, but on August 10th, when tweets collected amount to 18.800, exactly twice the volume reached on the 12th. This is the day with the highest volume of tweets in BW channel during the whole monitoring period. Similarly, in the WW channel, the 10th is the day when most tweets were retrieved, but August 12th is not far behind in the count (respectively 3.729 and 3.456 tweets). Despite this particular feature, we highlight that

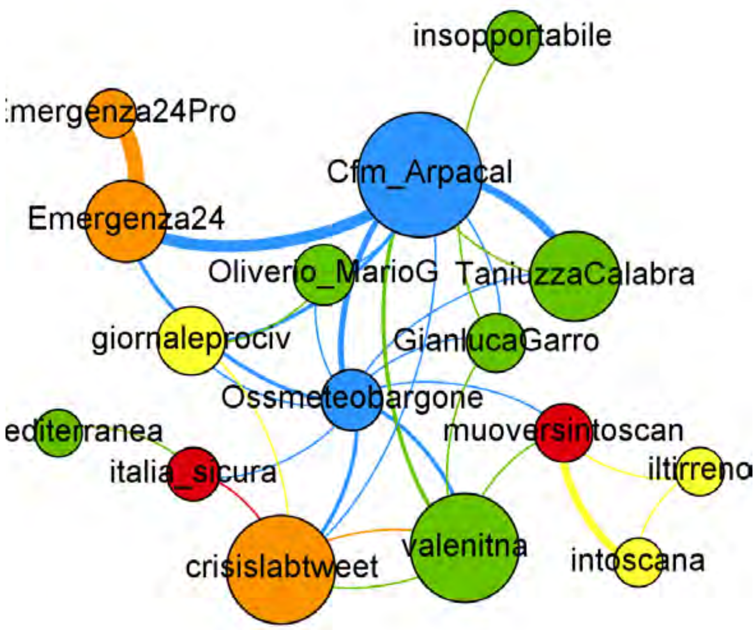


Figure 7.6: Retweet graph of codified hashtags during the Calabria flood. Node size is proportional to the value of betweenness centrality. Node color relates to users category: red for Institutions; green for Citizens; orange for NGO; yellow for Media; blue for Weather services.

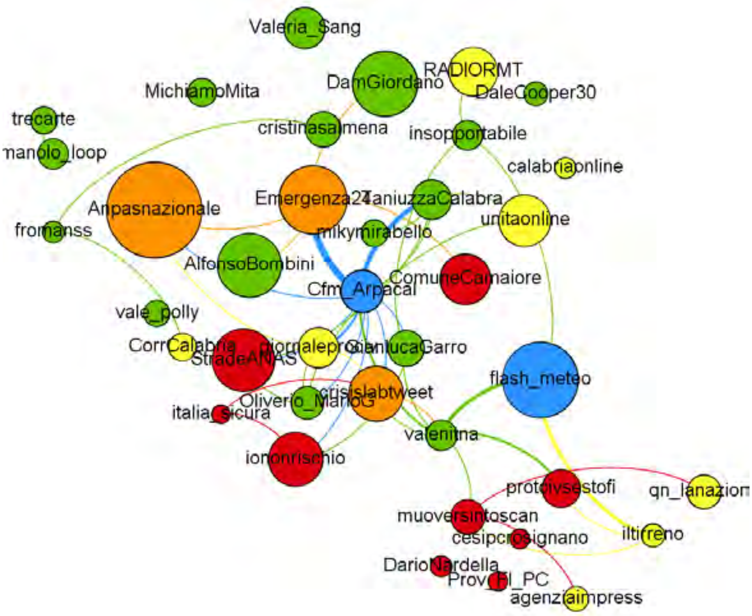


Figure 7.7: Retweet graph of codified hashtags during Calabria flood. Node size is proportional to value of eigenvector centrality. Node color relates to users category: red for Institutions; green for Citizens; orange for NGO; yellow for Media; blue for Weather services.

Table 7.6: Main features of the monitored channels on August 12th, during the Calabria flood.

Channels	**TW and RT**	**TW**	**RT**	**Unique users**
BW	9.441	5.897	3.544	6.016
WW	3.456	1.781	1.675	1.795
FU	627	139	488	388
CH	1.190	201	989	667

in both channels, August 12th results to be the day with the greater volume of retweets compared to the previous days (a 50% of RT compared to a 30% registered as average for the previous days). Differently on FU channel, volume pattern is similar to the Codified Hashtags channel with major tweets published on the Rossano flooding day and a retweets rate of 80% of the whole set of messages. As shown in table 7.6 not only the total volume of tweets is very different but also the number of unique users is so. On August 12th the number of unique users employing the codified hashtag to share information about Calabria floods or other weather related emergencies in Italy was around a tenth of people tweeting about bad weather conditions and a third of those sending tweets about #weather. The same proportion is valid for tweets volume.

7.4 Case 3: widespread rainfalls on the Italian peninsula

7.4.1 Meteorological scenario

At the end of February 2016, a strong wave of bad weather affected the whole Italian peninsula, starting from the Northern regions and moving to the Central and Southern part of Italy. On February 28th heavy rains fell over great part of Piedmont, Liguria, Tuscany, Lazio, accompanied by intense storms also on the coast. On Monday February 29th bad weather conditions mainly affected the Center-South of the country, with torrential rains and strong winds that blew down trees and caused overflowing of streams. The Central Functional Center of the Italian Civil Protection issued an orange warning for landslide risk in Lombardy and Abruzzo and for flooding and landslide risk in Tuscany, Umbria, Veneto, Emilia Romagna, Campania, Basilicata and Calabria. A yellow warning for the same risks was issued for most of the remaining Italian peninsula, as well as in the North of Sardinia. In Emilia Romagna, due to the large amounts of rainfalls, rivers and streams reached a high danger level; in the district of Piacenza the stream Chiavenna overflowed. The Panaro River in the city of Modena (Emilia Romagna) reached a very high warning level and also did Seveso and Lambro rivers in Lombardy. This case study was selected because it nearly involved the whole country and orange warnings were issued for half of the Italian regions, thus offering a good case to evaluate the effective use of codified hashtags in different contexts. We analyzed tweets published from February 28th to the 1st of March 2016.

7.4.2 Tweets analysis

Taking into account all tweets collected by the TwitterVigilance platform in the Codified Hashtags channel, we analyzed messages published between Feb. 28th to Mar. 1st. The collected tweets total up to 602 messages, with 203 native tweets and 399 retweets. Retweets are 66% of the whole set of collected messages. Tweets were published by 280 unique users, 75% of them contributing only by retweeting. The majority of tweets were mainly published on Feb. 28th and 29th, while users were slightly higher on Feb. 29th, when the retweets percentage compared to the daily volume of messages reached 91%. Compared to previous case studies this one is much smaller, but it was nevertheless interesting since many regions were involved giving the chance to compare codified hashtags adoption in different contexts. In this sense the analysis of the used **hashtags** is very interesting. Table 7.7 shows which hashtags occurred more frequently in the data set: #allertameteoER and #allertameteoTOS are the most used, followed by #allertameteoLIG and #allertameteoPIE. This event involved a great part of northern Italy so it is no surprise to find so many codified hashtags. Among the less used appear: #allertameteoSAR, 10 times; #allertameteoABR 3; #allertameteoBAS 2; #allertameteoCAL 2; #allertameteoLOM 2; #allertameteoVEN 2. Obviously a few regions like Lombardia and Veneto did mot use at all the codified hashtag on Twitter. Total **hyperlinks** retrieved in the data set of codified hashtags during the monitored days were 461, corresponding to 178 unique URLs. Considering only the 203 original tweets published during the three days, URLs were only 5. **Favorite** tweets are 143, 88% (126) of which

Table 7.7: Most used hashtags of selected events in CH channel

Top 10 hashtag	**freq**
#allertameteoer	294
#allertameteotos	169
#allertameteolig	94
#modena	50
#maltempo	31
#secchia	31
#allertameteopie	29
#bastiglia	23
#firenze	23
#idrogeologico	23

recorded on Aug. 12th.

Among the top twenty **unique authors** of native tweets, table 7.8 shows that main contributors are Institutions and Media accounts. The majority of them is from Emilia Romagna, while only a few are from Tuscany. Even if the most used hashtag is the one for Emilia Romagna, weather accounts included in the top 20 are not regionally relevant. One of them @ossmeteobargone is the account of the meteorological Observatory of Bargone Casarza, near Genova, Liguria, while another one is the account of Calabria Functional Centre. The reference account for weather in Emilia Romagna (@arpaER) is not listed in the top 20 users. The most **mentioned users** are distributed among the different regions impacted by the bad weather. Table 7.9 shows the 20 most mentioned users of codified hashtags channel during this event. Users are distributed among the different regions impacted by the bad weather conditions. The account of @arpaER is the seventh most mentioned, though the most mentioned weather account is @wwwmeteoit, a very famous commercial forecasting service. The most mentioned institutional accounts are from all the different regions impacted: Liguria (@LiguriaOnLine), Tuscany (@protcivileIolo and @ProtCivComuneFi), Emilia Romagna (@ProtCivGranarolo and @ComuneRubiera), Sardinia (@Comune_Cagliari) and Piedmont (@PCProvAL), meaning that when institutions are contributing on Twitter during emergency they are followed. The most mentioned media accounts are both from Emilia Romagna as consequence of the expected surge of small rivers near Modena.

Table 7.8: the top 20 authors of the codified hashtags channel during the selected event

Authors	**Category**	**Occurrance**
ossmeteobargone	weather	33
amtoscana	bot	26
gazzettamodena	media	21
fangareggi	citizens	17
procivgranarolo	insitutions	17
24emilia	media	16
capotopi	citizens	14
edii67d	citizens	14
radioperusia	media	11
bacci7_bacci	citizens	10
cfm_arpacal	weather	10
emergenzeprato	bot	10
vabprato	NGO	9
finaleemilia	insitutions	8
francescodondi	citizens	8
misepalazzi	NGO	7
tempo_di_carpi	media	7
itnewsfi	media	6
matcamme	insitutions	6
protciviolo	insitutions	5

Table 7.9: Most mentioned users of codified hashtag channel

Mentioned user	**Category**	**Frequency**
24emilia	media	41
cesipcrosignano	institutions	38
gazzettamodena	media	34
LiguriaOnLine	institutions	29
wwwmeteoit	weather	22
FrancescoDondi	citizens	16
ArpaER	weather	15
intoscana	media	14
ProtcivIolo	institutions	13
AllertaMeteoLIG	bot	12
ArpaPiemonte	weather	12
Miti_Vigliero	citizens	11
ProtCivComuneFi	institutions	10
iltirreno	media	10
ProCivGranarolo	institutions	9
quotidianopiem	media	8
Comune_Cagliari	institutions	7
ComuneRubiera	institutions	7
ConfesercentiMO	citizens	7
PCProvAL	institutions	6

7.4.3 Network analysis

Social Networks Analysis was used to identify important users (nodes) in the network. See section 4.3.3 for description of procedure. Figure 7.8 and 7.9 represent the retweet graphs of communication propagation around the codified hashtags during the Calabria floods. Nodes in the graph are unique users and the node sizes measure the Eigenvector or Betweenness centrality. Graphs are generated using Gephi software (an open-source software for network visualization and analysis); the network layout employed is *ForceAtlas 2*, with default settings. Given the high number of nodes, the graph in figure 7.9 shows only nodes with Eigenvector values ranging from 0.068 to 1.

In figure 7.8 the graph shows Betweenness centrality of the network, filtered by values ranging from 1.74 to 116.5. There's a clear partition of the graph: we may see that one part of the graph is more related to weather alert in Tuscany and Liguria, while the other one to weather alert in Emilia Romagna. Nodes with a hub role are Media, like @GazzettadiModena, and @24_Emilia and @OssmeteoBargone (a weather service in Liguria). As emerged also in mentions statistics, many institutional accounts are related to Tuscan Municipalities and Civil Protection.

As for the Betweenness graph, figure 7.9 highlights two branches expressing different groups of users, the ones below from Emilia Romagna and the ones above from Tuscany and Liguria. In the lower part of the group greater nodes are all green, meaning that Citizens are the most influential users; Institutions are rarely present. On the contrary, a copious amount of institutional accounts show up in the upper part of the graph, and many of them are from Tuscany. In Tuscany and Liguria institutions are the most central players of the codified hashtag community.

7.4.4 Codified hashtags diffusion

As previously mentioned, this case is smaller in terms of collected tweets compared to others considered in this work. It is interesting for the purpose of this study because it covers a period of days of bad weather conditions all over the Italian Peninsula, giving us a chance to compare the effective use of hashtags for different regions. We saw earlier that, despite many regions were impacted by bad weather conditions during the days between the end of February 2016 and the beginning of March 2016, only Tuscany, Emilia Romagna and Liguria used the hashtag in a consistent way. The volume of

Figure 7.8: the retweet graph of codified hashtags. Node size is proportional to the value of betweenness centrality. Node color relates to users category: red for Institutions; green for Citizens; orange for NGO; yellow for Media; blue for Weather services. Nodes have been filtered by values ranging from 1.74 to 116.5; nodes with lower values have been excluded.

Figure 7.9: The retweet graph of codified hashtags. Node size is proportional to the value of eigenvector centrality. Node color relates to users category: red for Institutions; green for Citizens; orange for NGO; yellow for Media; blue for Weather services. Nodes have been filtered by eigenvector centrality value ranging from 0.068 to 1; nodes with lower values have been excluded.

Table 7.10: Tweets collected by TwitterVigilance channels on Feb. 28th 2016.

	TW + RT	**TW**	**RT**	**RT %**	**Unique Users**
CH channel	292	161	131	45%	136
BW channel	15.517	9006	6511	42%	10.219
WW channel	3543	2326	1217	34%	1853
FU channel	541	180	361	67%	271

tweets has been, in any case, very low compared to other events examined in previous sections. To have a measure of the use of codified hashtags during the observed period, we compared tweets of CH channel with tweets collected from the more general channels regarding bad weather conditions (BW) or weather in general (WW), and those related to weather forecasting services (FU). In BW channel, the volume of collected tweets for the days Feb. 28th and 29th is more or less the same, around 15.500, with a proportion of retweets around 40% on the whole set. Unique users are around 10.000 on Feb. 28th and 10.200 on the 29th. Compared to them, unique users taking part in Twitter communication around codified hashtags were the 1%, just a drop in the sea of Twitter. Greater volumes of messages don't mean necessarily more information; within BW channel the most retweeted message was a tweet by a young singer sending a picture of him under the rain, that was retweeted 2132 times and got 3835 likes. Refining the retrieval only to tweets containing the #maltempo, we find that the most retweeted message was a lot less popular, with 22 retweets and 23 likes. It was a tweet from the media account of @MediasetTgcom24 (television news service) about a train derailment due to a landslide in Biella (Lombardy). Tweets collected in WW channel are fewer than those in the BW channel; anyway they are still 10 times those collected with codified hashtags channel, as for unique users. Retweets percentage is even lower, 34%, than BW channel. The last channel, FU, collecting messages mentioned by or published by weather forecasting services, shows almost twice the tweets and users of the CH channel.

7.5 Conclusions

In this chapter three events are analyzed in order to gain more insight about the effective use of codified hashtags for weather warning during critical situations. Besides Twitter metrics, coding of users has proved to be a useful tool to highlight the role that different categories of actors played in these emergencies.

The analysis enforced the outcome of the previous chapter: regions adopting the codified hashtag during the monitored period turned out to be Tuscany, Liguria, Emilia Romagna, Sardinia and Calabria; other important regions like Lombardy and Veneto, although interested by severe weather warnings did not show any substantial uptake. The analysis of users showed a significant occurrence of institutional accounts from Tuscany also among the most active users examined for the Sardinia and Calabria case studies. By modeling information diffusion with the help of Social Network Analysis we were able to show in a retweet graph how different categories of users played an influential role in the studied contexts. In Sardinia influencers in the hashtag-community were Citizens. In Calabria the most important users in the retweet graph belong to Media outlets, emergency management NGO, citizens and weather services, with lower occurrence of local institutional accounts and higher of national NGO or Weather Service. A similar pattern emerges also in the analysis of the third case study. Results showed how the adoption of the codified hashtag in certain contexts might be promoted also by non-local accounts of people or organizations that got involved remotely in the emergency. The digital volunteers willing to participate in the relief effort may find in the hashtag-community the virtual place where to offer help, also from remote locations. This confirms what disaster sociologists report about emergent organizations during disasters. These groups arise when the community feels that institutional structures and existing organizations do not meet information needs and new demands coming out from the disaster [8, 174, 197, 198]. Digital technologies highly enlarged the composition and size of these groups of volunteers, thus including also people in remote locations who offer skills and time for emergency response. These groups of users that gather around the hashtag don't know each other and don't even need to be connected by pre-existing follower networks. They are an example of community forming dynamically that is intrinsically "ephemeral". The codification of a hashtag for emergency response may contribute to make this community more "stable".

Furthermore, when institutions support the codified hashtag, like it happened in Tuscany and just started in Emilia-Romagna and Liguria, the formed hashtag-communities take the shape of emergent groups where formal organizations and volunteers find a common space of potential interaction. This doesn't mean that citizens' activities are automatically and easily integrated into emergency response procedures, but it assures that institutions and emergency managers might at least monitor what is being organized and might prevent potentially unsafe or risky actions during the response phase.

As reported by many studies, institutions still meet many problems in using social media during disasters, for organizational and procedural reasons. In closing this chapter we add a list of best practices that institutions could follow to improve their use of Twitter for emergency management. We organized the list by following the temporal phases of an emergency.

BEFORE the emergency

To organize an effective social media presence during an emergency is very important to do a lot of work before the disaster strikes.

- First of all, it is fundamental to have already created an online reliable identity. The profiles of institutional organizations should have a clear and recognizable identity. A link to the institutional website should always be included in the profile description and also a link to the social media policy. This is a document where the organization presents its online strategy and draws the operational guidelines for social media use in an emergency. Also specifying the actors and contents that can be delivered by the different social media in use. A list of the most used hashtags should be included, including the codified hashtag for weather warning or another accurately decided with the local community.

- It is important to set up pre-engaging activities on Twitter as to increase the institution credibility as an accurate and trusted source of information. Before the emergency, it is also critical to establish network connections with the media, local influencers and other institutions to ensure that this network will function and be coordinated during a crisis. By engaging local influencers before the emergency, institutions may count on a few trusted and high visible sources of information that during the crisis will receive and share important news and advice, diminishing the risk of misinformation. For what concerns the style of tweets, it is better not to rely on automated publishing systems that are easily recognizable by users. It is far better to manually write every post, using a warm style that gives to the organizations that human-touch important to engage the public. Mentions and retweets of trusted sources are an essential component of the engagement strategy.

- Another important element to implement in advance is the creation of a hub-website that conglomerates as many resources and information as possible in one place. This hub of information may then be linked to other official websites and shared through social media. In order to use the codified hashtag to aggregate all emergency- related information, it could be useful to include in the hub-website also a page with the Twitter widget to show and a hashtag timeline. In this way it would be possible to share with the general public the information produced on Twitter. (This hub website has been implemented during 2016 - 2017 by the regional government of Liguria and Emilia Romagna).

- Last but not least, it is very important to set a monitoring activity to analyze actors and features of the online communication about weather and emergencies. Monitoring Twitter to find out recurrent hashtags and main information sources. Monitoring is of course a core activity also during the emergency event.

DURING the emergency

During the disaster, it's important to assure an immediate, accurate and continuous presence online.

- In this phase monitoring Twitter is very important for several reasons. Firstly, to identify and engage with relevant users who can be an important and trusted source of information and also a powerful channel for sharing advice. Engaging actively with these users is critical to coordinate institutional communication with online volunteers and emergent groups activities. Furthermore, monitoring Twitter is important to find out inaccurate information or unsafe activities and fix them by giving proper recommendations. The monitoring of trending hashtags is also helpful to redirect messages and promote and sustain the appropriate codified hashtag. When the codified hashtag is used inappropriately, like for advertising or self-promotion, its use should be reproached. For hashtags monitoring is essential to follow as well the tags of locations interested by the emergency. Those are in fact very often cited as hashtags in tweets and are helpful to identify other emerging hashtags as the event unfolds.

- Monitoring is still helpful even when institutions or agencies are not yet using Twitter as an official communication channel. Twitter monitoring can still proves to be helpful in increasing situational awareness by easing the identification of trends or hitches in the communication process. This can be done with the support of specific software or algorithms that capture and automate messaging through machine-readable syntax in order to map the data for situational awareness.
- Institutions or governmental agencies should try not to merely use Twitter to disseminate updates and advice. They instead should use it as a two-way communication channel, possibly responding to requests, correcting misinformation, and debunking rumors and hoaxes. Activate as much as possible a conversational communication style in order to engage citizens and other volunteers and interact with them. This will increase the sense of community and the trust in official agencies. We know that this conversational style is still very difficult for institutional accounts, especially during crises, when officials are busy with emergency management activities and no social media staff is available.
- During weather related emergencies is very important to share and disseminate all the weather updates published by the official sources. When regional weather services have no Twitter accounts, other institutional users may share the updates using the codified hashtag.
- Verification of information is essential. Once information is verified it is worthy to share it.

AFTER the emergency

- Once the emergency is over, institutions should report on Twitter about the disaster activities and the planned reconstruction phases.
- In case of missing a warning or in case of a false alarm, it is important to explain why it happened in order to make people aware of the functioning of the alerting system and of the limits of weather predictability. Openness will contribute to increase trustworthiness.

- Hashtags monitoring in these phases is still precious to identify and reply to potential criticisms or incorrect information.

Chapter 8

Codified hashtag in a very localized event

In this chapter we present an analysis of the use of the codified hashtag #allertameteoTOS during a very local and extreme event, happened on August 1st 2015 in Florence. As Tuscany is the region where the codified hashtag for weather warning has been supported by the Institutions, analysis of this case helps to assess the use of it during a very particular event. For the analysis we retrieved the tweets containing the codified hashtag #allertameteoTOS and other relevant hashtags, like #Florence, #maltempo (bad weather) and #nubifragio (cloudburst). The data sets of collected tweets were compared in order to analyze main metrics. A manual annotation of users and tweets content was also performed with the aim of achieving a better understanding about the roles of different class of users and about the typology of information shared.

8.1 The weather event

On Saturday, August 1st 2015, in the late afternoon a strong storm with torrential rains, hail and violent gusts of wind dropped over Florence. The areas to the south-east of the city were fiercely hit by the storm with widespread flooding and a lot of damage due to falling of trees. On Friday July 31st the LaMMA weather service issued a yellow alert for thunderstorm risk on the entire region of Tuscany . In the city of Florence, the local meteorological conditions caused an exceptionally violent event, scientifically called downburst [1]. It caused a lot of damage in the Southern part of the city, with 1240 trees falling down or resulted dead afterwards. Many public buildings were damaged, many of them were schools. The storm injured 20 people and one has been seriously wounded, it blocked roads and caused damage to vehicles and motorcycles, and it resulted in various flooding.

8.2 Data analysis

To analyze the event we used a selection of tweets collected by different channels of the TwitterVigilance platform. The first channel to be considered was the Codified Hashtags channel. A further investigation covered other relevant hashtags related to bad weather conditions retrieved by the platform. The analysis covered the tweets collected from August 1st to August 3rd, 2015. The tweets collected were 660, composed by 111 tweets (17%) and 549 retweets (83%). Unique users that used the codified hashtag were 297, but as few as 26 if we consider only native tweets.

Looking at these small figures, it has to be considered that August 1st was part of the first week end of Summer holidays for many Florentines, so the city was for sure emptier than usual. Few people were engaged on Twitter because few people were in Florence. Furthermore, the event impacted just a confined part of the city in the Southern area, while other parts of the city only experienced a normal thunderstorm.

To better understand engagement of different category of users, unique users were coded following categories as identified in previous section 4.3.2. Authors belonging to Citizens were the most numerous, 42% of the unique users of native tweets; Media accounted for 27%, Institutions for 15%, NGO for 12% and Weather only for 4%. The small mentions of weather service accounts can be explained by the very localized nature of the event that did not allow to forecast the storm neither to monitor its evolutions as it unfolded. The accounts of Florence Municipality, both the official account of Comune di Firenze, @ComuneFi, the personal account of the Mayor Dario Nardella, @darionardella, and the account of the Florence Civil Protection,

[1]Downbursts are created by an area of significantly rain-cooled air that, after reaching ground level, spreads out in all directions producing strong winds. For this reason it may be confused with a tornado. https://en.wikipedia.org/wiki/Downburst

Table 8.1: Most used hashtags mentioned in CH channel during the event

Hashtag	**Occurrence**
#allertameteoTOS	654
#firenze	325
#maltempo	240
#nubifragio	235
#meteo	195
#allagamenti	79
#trenitalia	40
#treno	40
#protezionecivile	31
#anconella	13

@prov_fi_pc, were very active during the event. They used Twitter to post recommendations about road conditions and safety measures.

The down-burst hit Florence in the late afternoon of August 1st. Even if the storm only lasted one hour, it caused heavy damages in the area of the city that was impacted. Distribution of tweets in the monitored period is of 301 tweets during the first day, 345 during the second one and just 14 during the third one. Retweets were respectively 85%, 84% and 29%. Despite the small data set, the hashtags employed account for 1916 occurrences of 27 unique hashtags. As it usually happens during emergencies, one of the most used hashtags was #Firenze (Florence), name of the impacted location, followed soon afterwards by more generic words related to bad weather conditions, like #maltempo and #nubifragio (cloudburst). This last word was introduced by Media accounts which used it in titles and tweets to tag the event. This is another example of how the mass media shape greatly how the general public perceives and responds to hazards and disasters, also influencing individual disaster knowledge, attitudes, and behaviors [88, 158].

8.2.1 Different communities for different hashtags

To study the communication dynamic of this event, we compared Twitter conversations developed around the most used hashtags, as emerging by the TwitterVigilance data sets. Among the tweets collected by the platform, we filtered tweets for four different hashtags: 1. #allertameteoTOS; 2. #firenze;

Table 8.2: Volumes of tweets and retweets

	#amTOS	**#firenze**	**#maltempo**	**#nubifragio**
Native tweets	97	336	89	196
TW + RT	417	1055	723	534
Ratio RT/TW	3.3	2.1	7.1	1.7
01/08 TW	32	74	28	18
01/08 RT	182	314	29	161
02/08 TW	65	262	61	178
02/08 RT	235	741	694	373

3. #maltempo; 4. #nubifragio. Main metrics were calculated for each data set. We considered tweets retrieved by the TV platform between 7 PM of Aug. 1st to 11:59 PM of Aug. 2nd, and including at least one of the identified hashtags. Querying parameters for the selected hashtags were already active within the platform. Table 8.2 summarizes the main activity metrics. Tweets containing the hashtag #Florence were higher (1055) than others: they more than doubled those published using the codified hashtag for Tuscany, #allertameteoTOS (417), or #nubifragio (534), and were a thirty per cent more than those of #maltempo (723). As a matter of fact people tend to use, or include, as hashtag the name of the impacted location. In all data sets retweets were far more than tweets, as it is typical of disaster situations [177]. Data set with hashtag #maltempo shows the highest ratio among tweets and retweets, 7 retweets per native tweet, on average, while for #nubifragio ratio is only 1,7.

8.2.2 Temporal distribution of tweets

The down-burst impacted the southern area of Florence between 7 PM and 8 PM of August 1st (local time). First tweets were published around the same time using the hashtag #Firenze. Figure 8.1 shows temporal distribution of the hourly volume of tweets along the monitored. The first peak of Twitter activity is around 8 PM and 9 PM in the evening of August the 1st, soon afterwards the end of the violent down-burst. The second surge is in the morning of the following day, with many posts about road conditions and several reassurance messages published by Institutions and Civil Protection working in the field.

Some tweets published on the day before were about a yellow warning

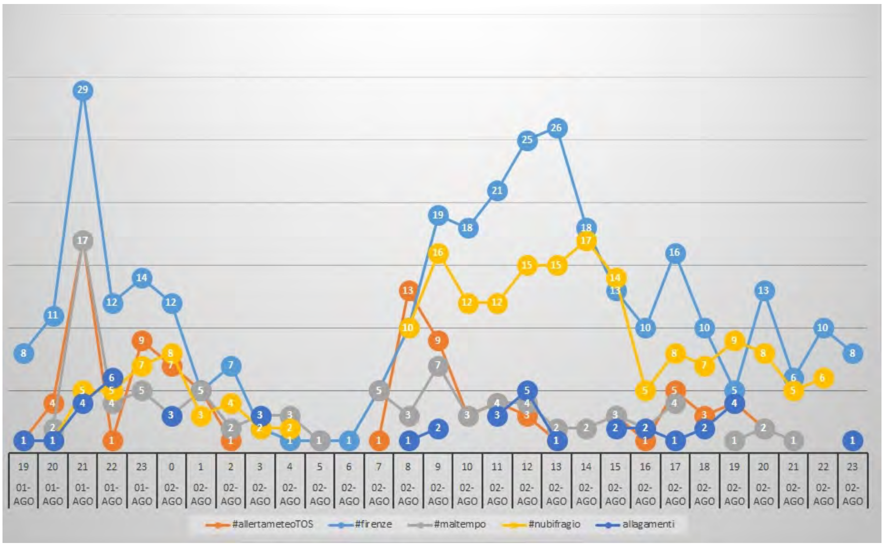


Figure 8.1: Hourly distribution of native tweets containing the 4 hashtags

for storm risk issued by the regional forecasting service, LaMMA. An example is this one posted by the account of civil protection in Rosignano @cesipcrosignano

> *"#allertameteotos codice giallo per temporali previsti per domani su tutta la regione http://t.co/c66c2bc0xp"*.

On August 1st, the first message to use #allertameteoTOS was posted at 8:13 PM reporting of some flooding in the surroundings of Florence. Around eight o'clock, a message from @radiofirenze reported of a storm and informed that a few basements had been flooded on the hills near Florence. The first tweet about the storm reporting of severe impacts in the city of Florence was published at 8:13 PM by @controradio, a well-known local radio station. It informs that a storm with strong wind gusts hit Florence. At 20:23 it is posted the first message mentioning a disaster. It is from a personal account (@framar1977) and it reports about *"the disaster of Florence, storm and dozens of felled trees"* sharing the hyper-link of the news just published on the website of the Florentine newspaper La Nazione. Actually this account belongs to the editor of the newspaper, who immediately shared the story and photos on Twitter. Looking at the hashtag #Firenze, we find several tweets that were sent during the storm. They were from citizens sharing pictures of lightning. A video about the storm in Florence was publshed around 7.22 PM. It was retweeted ten times and got 4 likes. It was posted by the account of @marziofattucchi, a journalist of the newspaper La Nazione. The first tweets posted soon after the storm impact are from local media outlet accounts, radio and news services. At 7.46 PM we find the first tweet reporting flooded streets:

> "*#firenze bomba d'acqua.allagamenti in zona via giampaolo orsini e coluccio salutati. problemi a #gavinana* " (#firenze water bomb. flooding in the area via giampaolo orsini and coluccio salutati. troubles in #gavinana).

The expression "bomba d'acqua" (water-bomb) started to being used some years ago after news reports titles in order to describe violent storms and flash-floods; it is still very used by common people and in the media, despite meteorologists and scientists strongly criticized the term.

Looking at #maltempo data set, the first tweet related to the Florentin event was posted at 8.42 PM by @retemeteoamator, an account of forecasting amateurs. It was the first message clearly assessing that the meteorological phenomenon was a down-burst and not an ordinary storm. Within #nubifragio data set, there are fewer tweets covering the first stages of the event, compared to #Firenze and #maltempo data sets. The hashtag entered into use lately, after the media employed the saying in news titles and online posts. By 10 PM in the evening of August the 1st., 25 tweets were posted using #allertameteoTOS, 33 using the hashtag #maltempo (badweather), 76 tweets with #Firenze and only 6 with hashtag #nubifragio (cloudburst). The most retweeted message was published at 9.59 PM by the official account of Florence municipality, @comunefi, using the codified hashtag #allertameteoTOS to prevent citizens from using the car until the following morning. It was retweeted 95 times and got 14 likes. The second most retweted message was published from the media account @MediasetTgcom24 with an image of a woman sitting on the bonnet of her car and talking on her cell phone against the backdrop of a flooded road:

> "*Temporale a Firenze. Cittá allagata. Le foto del disastro*" (Storm in Florence - flooded city. The photos of the disaster)

with a link to a photo gallery on the media outlet website. The tweet was sent at 10:10 PM and got 75 retweets and 40 likes and contained both the hashtags #firenze and #maltempo. As noted in previous works [175], during disasters users are more likely to retweet information originally distributed through Twitter accounts run by media, especially the local media, and traditional service organizations. Among the most retweeted messages, there are also the tweets from the Florentine mayor Dario Nardella, @darionardella. He was on vacation and followed the first moments of the emergency via Twitter. He wrote his first tweet on August 1st, at 11:23 PM (local time)

> "*anche se sono all'estero con famiglia partecipo via telefono a unitá di crisi @comunefi fra poco aggiornamenti #nubifragio #allertameteotos*" (even if I'm abroad with family I'm following closely via telephone @comunefi the crisis unit. Updates shortly #nubifragio #allertameteotos).

In the next sections we will discuss results about authors and contents of tweets.

8.2.3 Users coding and analysis

Another part of the work was to better understand which typologies of users were involved around the different hashtags conversation. We suppose that within #allertameteoTOS data set, we will find more institutional users than in other data sets. For the analysis of communication pattern of similar Twitter authors, unique users were classified in five categories (Institution, Media, Volunteers-NGO, Weather, Individuals), as mentioned earlier in section 4.3.2.

Figure 8.2 shows distribution of different categories of users in the four data sets; figures 8.3 and 8.4 present the different distribution of unique users of native tweets following their categories. Institutions is the more numerous category in #allertameteoTOS data set, (35% of native tweets unique authors) and in #nubifragio as well. Media accounts are well distributed in all data sets, meaning that they used all of the four hashtags to send tweets. Citizens are instead more on #nubifragio and #Firenze data sets, respectively %46 and 38% of native tweets authors. NGO users are more on #allertameteoTOS and #maltempo. Weather accounts are very poorly represented in the data sets because of the rapid and unpredictable nature of this kind of events. The storm was unpredictable in its localization and

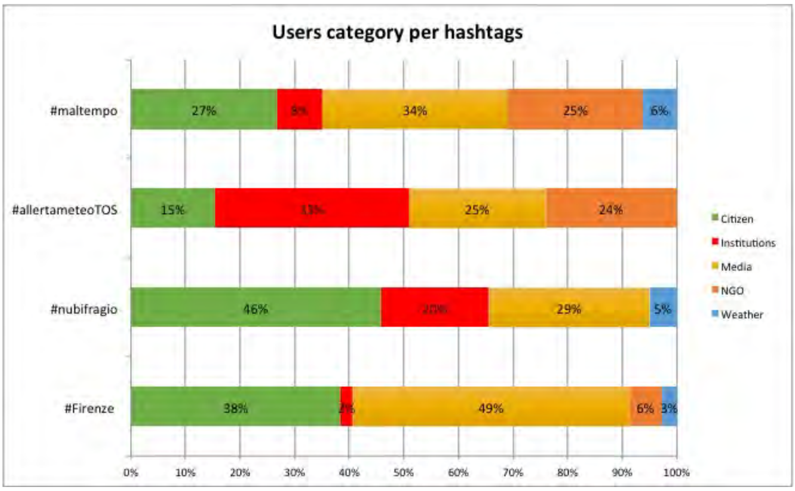


Figure 8.2: Users category publishing native tweets per different hashtags

impacts, thus Twitter conversations were more focused on management than on warning messages.

Local institutions were on the front line. @Comunefi published 24 tweets, 6 tweets between 9 PM of Aug. 1st and the early night and the rest the following day, August 2nd. The mayor @darionardella posted 8 tweets, a couple of hours after the event and the rest the following day.

8.2.4 Content coding and analysis

The analysis of users was important to assess how the different hashtags engaged different communities of users, particularly true for the #allertameteoTOS around which institutions in Tuscany communicate. Next step of the work was to analyze if hashtags also tend to define the nature of the messages shared. Hashtags function as conversation aggregators and if different tags gather different users they may also generate different class of contents. To perform the analysis, we examined messages published in the peak days of the event, August 1st and 2nd 2015, mentioning one of the monitored hashtags. Tweets were coded manually using selected categories already used in previous works [79]. We identified a set of eleven categories to describe the information communicated in each tweet, paying particular

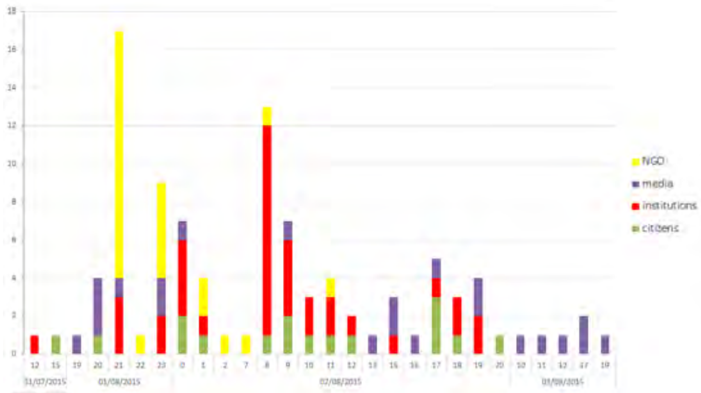


Figure 8.3: Native tweets per user category on #allertameteoTOS

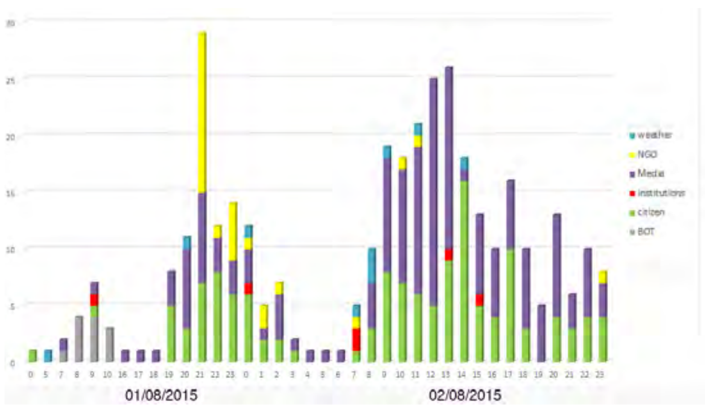


Figure 8.4: Native tweets per user category on #Firenze

attention to those that helped to increase situational awareness. Works by Starbird et al. (2010) [177] and Hughes (2014) [92] guided the identification of categories concerning situational update; two more categories were added to classify media contribution and comments shared by the public broadening the general understanding of emergency impact on the population. The categories considered were described in 4.3.2; they are Advice; Warnings; Hazard location/impact; Weather; Transport Conditions ; Evacuation and Closures; Damage; Reassurance; Resources; Comments; News reports. We also coded Off-topic messages.

Only native tweets were considered for content coding. Messages were distributed as follows: 336 for #Firenze, 106 for #allertameteoTOS, 117 for #maltempo and 317 for #nubifragio. Some differences are easily visible as summarized in table 8.3.

One-third of tweets of #Firenze data set belongs to news category. A 16% of posts were Comments and 13% were Hazard Location messages or tweets including Resources to get more information about the event; only 7% were Advice and 3% Evacuation information. Most retweeted messages of #Firenze data set are 2 tweets sent by a media news channel informing about the rescue of two people by the Fire Department. Messages with the hashtag #allertameteoTOS show a more informative nature with the majority of tweets delivering information about transport and road conditions (31%) and 20% publishing reassurance messages. This is related to the participation of institutional accounts, which employed the codified hashtags to connect with citizens during the emergency. 11% of messages are related to Damages reporting, delivered by citizens or by media accounts sharing pictures of the severe impacts in the southern part of Florence. In the codified hashtag data set, we also find a small bunch of messages about weather warning, before the event, and weather updates, in the aftermath. The most retweeted message is from Florence Municipality @comunfi (95 retweets) preventing citizens from using cars till the next morning. The second most retweeted message (33 retweets) is from the Florentine mayor @DarioNardella reassuring citizen that he was following the emergency though on vacation. A similar amount of retweets is obtained by a tweet from an NGO account, @emergenza24, informing about inundated roads and tunnels (32 retweets). The #nubifragio data set presents a different distribution of content types where the great part of native tweets are Comments sent by citizens, expressing sorrow and closeness to involved people. In this data set, contents which could be useful

Table 8.3: Messages by category

	#Firenze	**#amTOS**	**#maltempo**	**#nubifragio**
Advice	3%	6%	5%	0%
Comments	16%	1%	9%	28%
Damages	10%	11%	13%	12%
Evacuation	1%	2%	2%	2%
Hazard location	13%	6%	4%	5%
News	34%	13%	8%	15%
Off topic	1%	0%	18%	0%
Reassurance	3%	20%	4%	14%
Reassurance	13%	6%	14%	12%
Transport conditions	7%	31%	17%	9%
Warning	0%	2%	3%	0%
Weather	0%	3%	3%	3%

for situational awareness are less abundant. No Advice and Warning tweets were published using #nubifragio. The reason is that it was introduced by Media outlets in news titles and consequently adopted by people the following days.

Also #maltempo data set reveals to be a channel with less information abundance, not related to Tuscany and nor to this particular event. It is not surprising that the greatest content category is Off-topic (tweets not related to down-burst in Florence). The second and third more numerous category are tweets about Transport and Road conditions, Resources and those about reported Damages.

8.2.5 Network analysis

With the help of Gephi we modeled information propagation within users of #allertameteoTOS. Network graphs were realized for eigenvector centrality and betweenness centrality which are used to highlight influential users (for full description see sections 2.3.2 and 2.3.3). It's very clear that institutions played the leading roles. Figure 8.6 shows the retweet graph for eigenvector centrality, realized with Gephi by using ForceAtlas2 layout with default

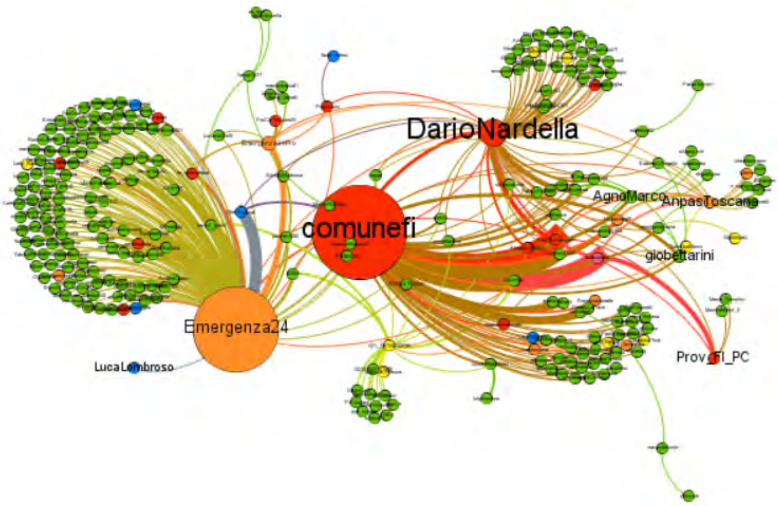


Figure 8.5: Retweet graph of the codified hashtag. Node size is proportional to values of betweenness centrality. Node color relates to users category: red for Institutions; green for Citizens; orange for NGO; yellow for Media; blue for Weather services.

settings. Colors stand for different users categories. Bigger nodes, those with higher eigenvector centrality, are the most influential users and are the account of Florence Municipality, @Comunefi, and the account of the Mayor of Florence, @DarioNardella. Nardella used its personal account to send reassuring messages to citizens, also preventing and responding to possible blaming for its absence (he was on vacation during the event). Figure 8.5 integrates these findings by showing nodes with higher betweenneess, those nodes that in the network functions as important bridges for information propagation, connecting different part of the network. This role is played by Florence Municipality @Comunefi, and @Emergenza24, an account of an emergency management association very active on Twitter. It's clearly visible that all other nodes are connected to them.

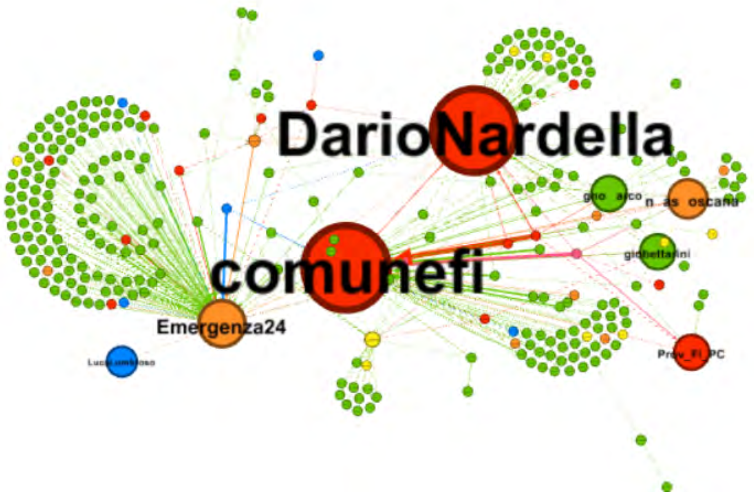


Figure 8.6: Retweet graph of the codified hashtag. Node size is proportional to values of eigenvector centrality. Node color relates to users category: red for Institutions; green for Citizens; orange for NGO; yellow for Media; blue for Weather services.

8.3 Conclusions

The analysis of this case study wanted to investigate the use of the codified hashtag in a case of a very localized and unpredictable event such it was the down-burst that hit the southern part of Florence on August 1st, 2015.

On this purpose, a selection of tweets stored by the Twitter Vigilance platform was analyzed comparing the messages using four different hashtags: #allertameteoTOS, #Firenze; #maltempo and #nubifragio. Authors were coded to compare the different categories of users involved in the conversations. Also, the content of native tweets was coded manually to fit into those categories identified as relevant in terms of their contribution to situational awareness. Analysis showed that the codified hashtag #allertameteoTOS was fairly used during the emergency. Even if tweets were less than those published with more generic tags, like the city name #Florence or weather related terms like #maltempo or #nubifragio. Institutions, NGOs, and media employed the codified hashtag to share information during the event. Citizens resulted in being the most numerous category on #Firenze and #nubifragio data set, while Media were visible in all the four data sets. Looking at the content of tweets, on #allertameteTOS most of the tweets were dealing with Road Conditions (31%) and with Reassurance messages (20%) sent by the Florence mayor account or by the Municipality account (31%). Institutions used the codified hashtag soon after the event to advise citizens and to reassure them in the following day. The Mayor, on vacation abroad, also used Twitter as the fastest way to reach population. A classic "personalization" of institutional communication that is somehow typical of the new digital era of public communication [121, 169], but that also contributes to information spreading (see Hurricane Sandy case discussion in section **??**.

Institutional accounts have a central role in the codified hashtag data set. This is confirmed by the provided network graphs, where the Twitter account of the Mayor of Florence, @darioNardella, and the official Florence Municipality account, @comunefi, hold an outstanding position.

Chapter 9

Conclusions

This chapter summarizes the contribution of the thesis and discusses perspectives for future research.

9.1 Summary of contributions

In recent years codified hashtagging has been proposed by some governments and disaster response organizations as a tool to improve communication on social media during disasters. Hashtags, in facts, are important to coordinate conversations on Twitter and to facilitate the retrieval of messages about a topic or an event. This research presented an analysis of the proposal of using codified hashtags for weather warnings in Italy on Twitter. For this purpose, we realized a one-year monitoring of twenty codified hashtags proposed for Italy by using a platform developed by the University of Florence to retrieve and store collected tweets. In this work, we presented an analysis of the whole set of hashtags for main Twitter metrics (chapter 6). We also analyzed three severe weather events to better understand the communication exchange and the hashtag dynamics during critical situations (chapter 7). Besides fundamental Twitter metrics of the hashtags data set, we also coded tweets for main categories of relevant users for emergency management. The retweet dynamics for the selected events were modeled using Social Network Analysis. For two case studies (chapter 5 and 8) we also performed a content analysis by coding the content of tweets for eleven categories that are recognized as relevant to situational awareness.

The analysis was carried out during a one-year time monitoring period of (July 2015 - June 2016). Results showed that codified hashtags have been used in Italy in about one-third of the regions, with a great diversity in the way they were adopted. Tuscany appears to be the only case of "*regular use*", as confirmed by the highest percentage of activity rate (69%), the great number of collected tweets and the highest ratio of tweets per user (12). In Tuscany, the codified hashtag was actually adopted and institutions had a primary role in creating a hashtag-community. Sardinia and Calabria show a "*burst use*" related to exceptional and isolated situations associated with the occurrence of a disaster. These data sets show the highest numbers of users engaged and the highest number of tweets, but only during the limited days of the emergency. The activity rate is around 25-30% during the whole monitored period. In these cases, Twitter activity around the hashtag is pushed by extraordinary circumstances and is largely sustained by Citizens. These are also the contexts where Institutions appear less engaged in the hashtag uptake. Emilia Romagna, Liguria and Sicily show a "*tentative use*". In Liguria and Emilia Romagna, the activity rate is about 35% indicating that the community is not yet fully engaged; temporal distribution of messages is more regular compared to Sardinia and Calabria. Institutions are not fully engaged in this adoption, but they have started to use the hashtag in tweets. In Sicily adoption appears weaker, with only 11% of activity rate. Piedmont is on a borderline. Its data set was not fully analyzed in this work because of the small amount of retrieved tweets, probably due to the limited number of severe weather warnings issued during the monitored period. Anyhow, in Piedmont a few institutions supported the codified hashtag. An evidence for this comes from data (published in Appendix A) showing that within the #allertameteoPIE data set the most retweeted user is @ArpaPiemonte, the account of the regional weather functional center.

In other contexts, codified hashtags seem poorly used during the considered period. Regions like Basilicata and Lazio seem to be at an "*embryonic stage*", with a limited use proposed by some authors but not regularly embraced by the community. The hashtag uptake was more structured and stable in contexts where the regional weather forecasting service, public administrations and local organizations in charge of emergency management adopted it, like in Tuscany. The central role of the regional weather service is proved by the fact that in Tuscany the Twitter account @flash_meteo was one of the most retweeted with 1265 retweets (see appendix A, table A.1). In Calabria as well the Twitter account of the Decentralized Functional Center @cfm_arpacal was the most retweeted with 822 retweets. The dissimilar hashtag uptake in the two contexts might be a consequence of the different size of their followers' network: the Tuscany weather service counts on more than 14.000 followers while the Calabria weather service on around 600.

When there is not such an institutional support, the hashtag will only go viral if "Twitter stars" (accounts with many thousands of followers) will use it, as in the case of Sardinia. As reported in Appendix A, table A.2, the most retweeted users in Sardinia data set are all citizens. This is perfectly in line with disaster sociology studies about convergent behaviors when they assess that emergent groups arise in disasters when the community feels that institutional structures do not meet information and management needs coming out from the disaster [8, 174, 197, 198].

In Tuscany, where the hashtag adoption was more regular and sustained by the engaged institutional accounts, codified hashtagging allowed the building over time of a **hashtag-based community** that has now become a sustainable virtual community of trusted users. This is proved by the fact that even if Tuscany did not see any major disaster during the considered period, the codified hashtag was still used and messages retweeted. When users are retweeting each other's messages, even in absence of a major emergency, this indicates that users are not just tweeting around the hashtag topic but are also following and reposting what others are publishing. The more this happens the more the hashtag-community may be said to act as a community [24]. Virtual communities open new possibilities for individuals to generate value and knowledge and to share it with others [10]. They may positively influence the enhancement of social capital and civic engagement. Nevertheless, the creation of a virtual space does not guarantee the sustainability of the community. This depends on the users willingness to continue sharing and exchanging knowledge, like it happens in stable social media communities about certain topics. During crisis, the use of multiple hashtags spontaneously emerging during an event do not lead to the same degree of knowledge sharing and hamper the continuance of the community. Social media have proved their use for emergency management but also have potential in building community disaster resilience. By using certain social media strategies, like hashtags codification, it is possible to support the creation of "communities of practice" before, during and after an event. Following Wenger definition, "communities of practice are groups of people who share a concern or passion for something they do and learn how to do it better as they interact regularly" [199]. The regular interaction of the members of the hashtag-community produce knowledge resources that may affect their practices during emergencies.

In relation to disaster sociology studies and the need to include emergent communities within disaster management procedures, the use of a codified hashtag when supported by institutional users could be an effective way to interconnect digital volunteers, institutional agencies and the general public. With new technologies and social media spreading, the involvement of citizens in disaster response is in fact inevitable. Therefore, it is important to find practical solutions to interconnect these activities and to manage the volunteers' participation. Building a dynamical community of practice around the codified hashtag could lead to include emergent online communities within disaster management procedures. As Perry assess, effective emergency planning should be "with" and not just "for" citizens [145]. When codified hashtag is adopted regularly, users may share information and build relationships that enable them to learn from each other. The hashtag community may collectively develops a shared repository of resources, which may include information, experiences, data, or practical solutions of addressing recurring problems.

The use of a codified hashtag in case of weather warning could be an effective strategy for this issue. The analysis showed that when institutions support the codified hashtag, as it happened in Tuscany and partially in Emilia Romagna and Liguria, the formed hashtag-communities become a kind of emergent group where formal organizations and volunteers may find a common space of potential interaction. This does not mean that citizens' activities are automatically integrated into emergency response procedures. But it assures that institutions and emergency managers may at least monitor via Twitter what is being organized and may prevent potentially unsafe or risky actions during the response phase. Therefore, it is important for institutions to build a trustworthy presence on Twitter and promote the codified hashtag before, during and after the disaster strikes. Some suggestions have been proposed in the conclusions of Chapter 7. Codified hashtag, therefore, may represent the first step to build a potentially more resilient local community characterized by a higher level of shared knowledge and practices among trusted sources.

Codified hashtagging demonstrated also to be an effective way to convey relevant information reducing the amount of useless or off-topic messages during the emergency. As shown for those case studies where we performed content analysis (chapter 5 and 8), the hashtag adoption produced a high rate of tweets related to situational awareness (up to 90% in a few cases) unlike other studies ([110, 193]).

Another contribution of this work is to have structured a monitoring channel of codified hashtags that led to a data set of more than 47.000 tweets [1] (and still growing). The use of a monitoring platform, like the TwitterVigilance, to retrieve and store tweets was very useful for the creation of a virtual *knowledge bank* [10] for future reference on Twitter's role during weather related emergencies. Thus, the data set of codified hashtags may be the ground for further investigations about engaged users and categories of contents shared within these virtual communities.

Even if social media research structured as a case study analysis does not allow to have results that are comparable with different situations, the use of quantitative case study based research has the potential to continue to help stakeholders to understand public interaction with social media during times of crisis in specific contexts, and therefore, the value of social media for crisis management.

9.2 Directions for future work

A few results of this work may lead to further investigations.

The manual annotation of the most active users turned out to be very useful to identify patterns of communications during critical days. By further expanding annotation to the whole set of users regularly engaged in the hashtag community, we could realize lists of trusted users to be implemented in platforms like the TwitterVigilance. This will allow the development of new type of metrics based on the engagement of different class of users that could be performed in almost real time.

Another possible direction of future work relates to the implementation of automatic recommendations systems for hashtags. The use of a handy platform like the TwitterVigilance, resulted in being strategic in this analysis. In particular, the possibility to easily set up different monitoring channels for tweets retrieval and storage is very important to monitor hashtags use in emergencies. The whole set of monitoring channels, structured with search terms semantically closed but with a different lexical specificity (from

[1] as reported by TwitterVigilance platform at the end of November 2016

#weather to #allertameteoXXX) allowed to have multiple listening channels to detect critical events. Critical situations in fact were easily recognizable on the TwitterVigilance for a similar visual pattern characterized by the synchronization of the relative maximum of daily tweets. By implementing specific functions in such platforms it could be possible to use statistics on keywords and trending hashtags of the different monitoring channels as a basis to implement automatic systems for hashtag recommendation in different users groups.

To better understand how in certain contexts the codified hashtag was adopted and in others was not adopted at all, it will be important to develop a qualitative social research directed to communication officers of different public administrations and organizations involved in the issuing and management of weather alerting in the different contexts. In many public organizations, in fact, social media use is strictly linked to communication procedures. For any further development it's important that organizational structures of public administrations fully understand this background. Technological solutions will not be able to overcome well structured "bureaucratic" practices without management awareness.

Appendix A

Appendix

This appendix presents some tables related to data discussed in Chapter 6, section 6.3.4. In particular the following tables present the data about the most retweeted users in the monitored period within the six codified hashtags data sets, with the amount of received retweets per authors and the coded category of each authors.

Table A.1: Most retweeted authors of #allertameteoTOS channel.

Tuscany		
Authors	**Category**	**RT count**
amtoscana	BOT	3497
Ossmeteobargone	Weather	1462
flash_meteo	Weather	1265
ProtCivComuneFi	Institution	1038
MisePalazzi	NGO	954
iltirreno	Media	679
Emergenza24	NGO	489
comunefi	Institution	451
cesipcrosignano	Institution	412
muoversintoscan	Institution	368

Table A.2: Most retweeted authors of #allertameteoSAR channel.

Sardinia		
Authors	**Category**	**RT count**
insopportabile	Citizens	1565
paoloigna1	Citizens	1392
llimantul	Citizens	1169
egyzia	Citizens	833
Ecatetriformis	Citizens	819
EddiePepsi	Citizens	784
AlienaKent	Citizens	782
PaoloMastino	Citizens	719
simonacambarau	Citizens	706
robertore62	Citizens	662

Table A.3: Most retweeted authors of #allertameteoCAL channel.

Calabria		
Authors	**Category**	**RT count**
Cfm_Arpacal	Weather	822
Ossmeteobargone	Weather	661
TaniuzzaCalabra	Citizens	575
Emergenza24	NGO	530
Oliverio_MarioG	Institution	507
michelelizzi	Citizens	446
paoloigna1	Citizens	370
GabrielePetrone	Citizens	361
daniereds1976	Citizens	335
000120o	Citizens	317

Table A.4: Most retweeted authors of #allertameteoLIG channel.

Liguria		
Authors	**Category**	**RT count**
RadioPerusia	NGO	1600
Ossmeteobargone	Weather	337
Emergenza24	NGO	330
ComunediGenova	Institution	308
liguriaonline	Institution	301
ansa_liguria	Media	297
AgorLerici	Media	295
AllertaMeteoIT	Weather	292
genesdegenes	Citizens	265
AllertaMeteoLIG	Weather	248

Table A.5: Most retweeted authors of #allertameteoER channel.

Emilia-Romagna		
Authors	**Category**	**RT count**
RegioneER	Institution	258
coscienzavigile	Citizens	240
Emergenza24	NGO	202
gazzettaReggioE	Media	187
Ossmeteobargone	Weather	187
bacci7_bacci	Citizens	150
24emilia	Media	141
ProvinciadiRE	Institution	133
pierbaia3	Citizens	112
CapitanAchab	Citizens	109

Table A.6: Most retweeted authors of #allertameteoSIC channel.

Sicily		
Authors	**Category**	**RT count**
Emergenza24	NGO	265
Ossmeteobargone	Weather	259
donatellaluisa	Citizens	249
DrpcSicilia	Institution	190
VabSicilia	NGO	186
Drpc_Informa	Institution	179
calogero_foti	Citizens	174
TuriCaggegi	Citizens	165
GraziaTardo	Citizens	164
giornaleprociv	Media	146

Table A.7: Most retweeted authors of #allertameteoPIE channel.

Piedmont		
Authors	**Category**	**RT count**
ArpaPiemonte	Weather	235
Ossmeteobargone	Weather	190
patrizia6119	Citizens	136
PiemonteInforma	Institution	85
Davide_on_a_Mac	Citizens	84
QCanavese	Media	83
Aringoogle	Citizens	68
5Tlive	Citizens	58
PCProvAL	Institution	58
PiemontAmbiente	Institution	49

Appendix B

Publications

This research activity has led to several publications in international journals and conferences. These are summarized below.

International Journals

1. A. Crisci, **V. Grasso**, P. Nesi, G. Pantaleo, I. Paoli, I. Zaza. "Predicting TV programme audience by using twitter based metrics", *Multimedia Tools and Applications*, iss. e2241v1, 2017.
   [DOI:10.1007/s11042-017-4880-x]

2. **V. Grasso**, A. Crisci, P. Nesi, G. Pantaleo, I. Zaza, B. Gozzini. "Italian codified hashtags for weather warning on Twitter. Who is really using them?", in *Advances in Science and Research*, 14, 63-69, 2017
   [DOI:10.5194/asr-14-63-2017 6]

3. **V. Grasso**, A. Crisci, M. Morabito, P. Nesi, G. Pantaleo. "Public crowdsensing of heat waves by social media data", *Advances in Science and Research*, 1, 1-10, 2017
   [DOI: 10.5194/asr-1-1-2017]

4. **V. Grasso**, A. Crisci. "Codified hashtags for weather warning on Twitter: an Italian case study", *PLOS Current Disasters* , 2016 July 5, Edition 1. [DOI:10.1371/currents.dis.967e71514ecb92402eca3bdc9b789529]

5. **V. Grasso**, I. Zaza, F. Zabini, G. Pantaleo, P. Nesi, A. Crisci. "Weather events identification in social media streams: tools to detect their evidence in Twitter", *PeerJ Preprints*, vol.4, iss. e2241v1, 2016.
   [DOI: 10.7287/peerj.preprints.2241v1]

6. **V. Grasso**, A. Crisci, F. Camilli. "Social trekking on Facebook: textual analysis and quantitative metrics for the promotion of tourism", *Agriregionieuropa*, vol. 10, iss. 36, 2014.

International Conferences and Workshops

1. **V. Grasso**, A. Crisci, P. Nesi, G. Pantaleo, I. Zaza, B. Gozzini. "Italian codified hashtags for weather warning on Twitter. Who is really using them?", *16th EMS Annual Meeting and 11th European Conference on Applied Climatology (ECAC)*, Trieste (Italy), September 2016.

2. **V. Grasso**, A. Crisci, M. Morabito, P. Nesi, G. Pantaleo. "Public crowdsensing of heat-waves by social media data", *16th EMS Annual Meeting and 11th European Conference on Applied Climatology (ECAC)*, Trieste (Italy), September 2016.

3. **V. Grasso**, A. Crisci. "Social media and severe weather events: mapping the impact footprint", in *Proc. of Social Media and Semantic technologies in Emergencies Response (SMERST)*, Coventry (UK), 2013.

4. **V. Grasso**, A. Crisci, F.P. Vaccari. "Evaluating communication efforts and urban climate constraints for energy savings. The Races project", in *Conference Proceedings: Two hundreds years of urban meteorology in the heart of Florence*, Florence (Italy), 2014.

5. **V. Grasso**, V. Capecchi, G. Bartolini, R. Benedetti, G. Betti, R. Magno, A. Orlandi, F. Piani, C. Tei, T. Torrigiani, F. Zabini. "Who want to be a weather forecast? Education and public outreach at LaMMA Consortium, home of Tuscany weather service", in *Book of Abstract. Science education and guidance in schools: the way forward*, Florence (Italy), 2013.

6. **V. Grasso**, F. De Chiara, M. Napolitano, A. Matese. "ArdOmino: let the sensors speak! Sharing local data on social media", in *The 2013 Open Reader - Stories and articles inspired by OKCon2013: Open Data, Broad, Deep, Connected (OkCon)*, Geneva (Switzerland), 2013.

National Conferences

1. A. Cavaliere, A. Crisci, **V. Grasso**, S. Menabeni, P. Nesi, G. Pantaleo. "Monitoring Public Attention on Environment Issues with Twitter Vigilance", in *CINI Annual Workshop on ICT for Smart City and Communities*, Palermo (PA), Italy, 2015.

Technical Reports

1. M. Morabito, A. Crisci, S. Orlandini, G. Brandani, A. Messeri, **V. Grasso**. "Report on Technologies Use and Cultural Factors", Deliverable of the Project CARISMAND, Cultural Factor and Risk Management in man-made and Natural Disasters (H2020-DRS-2014), Technical Report, 2016.

Bibliography

[1] D. Alexander, "Towards the development of a standard in emergency planning," *Disaster Prevention and Management: An International Journal*, vol. 14, no. 2, pp. 158–175, 2005.

[2] G. Arata, "Facebookpa," 2016. [Online]. Available: http://nexa.polito.it/nexafiles/FacebookPA_2013.pdf

[3] N. Arceneaux and A. Schmitz Weiss, "Seems stupid until you try it: press coverage of twitter, 2006-9," *New Media & Society*, vol. 12, no. 8, pp. 1262–1279, 2010.

[4] ARPA, "Rapporto tecnico evento 3-6 novembre 2014," Agenzia Regionale Protezione Ambiente, Roma, Tech. Rep., 11 2014.

[5] ——, "Rapporto tecnico evento 9-17 novembre 2014," Agenzia Regionale Protezione Ambiente, Roma, Tech. Rep., 11 2014.

[6] ARPAL, "Rapporto evento alluvione genova," ARPAL Centro Funzionale meteo-idrologico di Protezione Civile della Regione Liguria, Genova, Tech. Rep., 11 2014.

[7] Audiweb, "Sintesi dati della total digital audience di giugno 2015," Audiweb and Nielsen, Milano, Italy, Tech. Rep., June 2015, retrieved on November 10, 2016. [Online]. Available: http://www.audiweb.it/wp-content/uploads/2015/08/Report_dati_Audiweb_Giugno2015.pdf

[8] E. Auf der Heide, "Disaster response: principles of preparation and coordination," in *Disaster response: Principles of preparation and coordination.* Canadá. CV Mosby Company, 1989.

[9] L. Austin, B. Fisher Liu, and Y. Jin, "How audiences seek out crisis information: Exploring the social-mediated crisis communication model," *Journal of Applied Communication Research*, vol. 40, no. 2, pp. 188–207, 2012.

[10] S. Bandyopadhyay, S. Batabyal, S. Banerjee, and R. Malik, "Measuring the influence of dynamically created virtual communities anchored around hashtags," in *Business Analytics and Intelligence Conference.* Bangalore, India, 2015.

[11] A. H. Barton, "Social organization under stress: A sociological review of disaster studies." 1963.

[12] F. L. Bates and W. G. Peacock, *Living conditions, disasters and development: An approach to cross-cultural comparisons.* University of Georgia Press, 2008.

[13] K. J. Bennett, J. M. Olsen, S. Harris, S. Mekaru, A. A. Livinski, and J. S. Brownstein, "The perfect storm of information: combining traditional and non-traditional data sources for public health situational awareness during hurricane response," *PLOS Currents Disasters*, 2013.

[14] G. Blank and B. C. Reisdorf, "The participatory web: A user perspective on web 2.0," *Information, Communication & Society*, vol. 15, no. 4, pp. 537–554, 2012.

[15] F. Boersma, J. Ferguson, P. Groenewegen, and J. Wolbers, "Beyond the myth of control: toward network switching in disaster management," in *Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management ISCRAM*, 2014, pp. 125–129.

[16] A. Boin, "From crisis to disaster: towards an integrative perspective," *What is a disaster*, pp. 153–172, 2005.

[17] J. Bonnan-White, J. Shulman, and A. Bielecke, "Snow tweets: emergency information dissemination in a us county during 2014 winter storms," *PLoS currents*, vol. 6, 2014.

[18] D. S. Bortree and T. Seltzer, "Dialogic strategies and outcomes: An analysis of environmental advocacy groups' facebook profiles," *Public Relations Review*, vol. 35, no. 3, pp. 317–319, 2009.

[19] D. Boyd and K. Crawford, "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon," *Information, communication & society*, vol. 15, no. 5, pp. 662–679, 2012.

[20] D. Boyd and N. Ellison, "Social network sites: definition, history, and scholarship," *IEEE Engineering Management Review*, vol. 3, no. 38, pp. 16–31, 2010.

[21] D. M. Boyd and N. B. Ellison, "Social network sites: Definition, history, and scholarship," *Journal of computer-mediated communication*, vol. 13, no. 1, pp. 210–230, 2008.

[22] A. Bruns and J. Burgess, "Notes towards the scientific study of public communication on twitter," *Science and the Internet*, pp. 159–169, 2012.

[23] ——, "Crisis communication in natural disasters: The queensland floods and christchurch earthquakes," *Twitter and society*, vol. 89, pp. 373–384, 2014.

[24] ——, "Twitter hashtags from ad hoc to calculated publics," *Hashtag publics: The power and politics of discursive networks*, pp. 13–28, 2015.

[25] A. Bruns, J. Burgess *et al.*, "#ausvotes: How twitter covered the 2010 australian federal election," *Communication, Politics & Culture*, vol. 44, no. 2, p. 37, 2011.

[26] A. Bruns, J. E. Burgess, K. Crawford, and F. Shaw, "#qldfloods and @qpsmedia: Crisis communication on twitter in the 2011 south east queensland floods," Tech. Rep., 2012, available at: http://www.cci.edu.au/floodsreport.pdf (14 Jun. 2017).

[27] A. Bruns, T. Highfield, and R. A. Lind, "Blogs, twitter, and breaking news: The produsage of citizen journalism," *Produsing theory in a digital world: The intersection of audiences and production in contemporary theory*, vol. 80, no. 2012, pp. 15–32, 2012.

[28] A. Bruns and H. Moe, "Structural layers of communication on twitter," in *Twitter and society.* Peter Lang, 2014, vol. 89, pp. 15–28.

[29] A. Bruns and S. Stieglitz, "Quantitative approaches to comparing communication patterns on twitter," *Journal of Technology in Human Services*, vol. 30, no. 3-4, pp. 160–185, 2012.

[30] ——, "Towards more systematic twitter analysis: Metrics for tweeting activities," *International Journal of Social Research Methodology*, vol. 16, no. 2, pp. 91–108, 2013.

[31] ——, "Metrics for understanding communication on twitter," *Twitter and society*, vol. 89, pp. 69–82, 2014.

[32] I. Burton and R. W. Kates, "The floodplain and the seashore: a comparative analysis of hazard-zone occupance," *Geographical review*, vol. 54, no. 3, pp. 366–385, 1964.

[33] I. Burton, R. W. Kates, and G. F. White, "The environment as hazard," 1978.

[34] S. H. Burton, K. W. Tanner, C. G. Giraud-Carrier, J. H. West, and M. D. Barnes, ""right time, right place" health communication on twitter: value and accuracy of location information," *Journal of medical Internet research*, vol. 14, no. 6, p. e156, 2012.

[35] D. Buscaldi and I. Hernandez-Farias, "Sentiment analysis on microblogs for natural disasters management: a study on the 2014 genoa floodings," in *Proceedings of the 24th International Conference on World Wide Web.* ACM, 2015, pp. 1185–1188.

[36] C. Capineri, C. Calvino, and A. Romano, “Citizens and institutions as information prosumers-the case study of italian municipalities on twitter,” *International Journal of Spatial Data Infrastructures Research*, vol. 10, pp. 1–26, 2015.

[37] L. Carter, J. B. Thatcher, and R. Wright, “Social media and emergency management: Exploring state and local tweets,” in *System Sciences (HICSS), 2014 47th Hawaii International Conference on*. IEEE, 2014, pp. 1968–1977.

[38] H.-C. Chang, “A new perspective on twitter hashtag use: Diffusion of innovation theory,” *Proceedings of the American Society for Information Science and Technology*, vol. 47, no. 1, pp. 1–4, 2010.

[39] W. Chen and D. Sui, *Influenza surveillance using volunteered geographic information (VGI): A GIS-based Hidden Markov modeling approach.* publisher not identified, 2010.

[40] F. Cheong and C. Cheong, “Social media data mining: A social network analysis of tweets during the australian 2010-2011 floods,” in *15th Pacific Asia Conference on Information Systems (PACIS)*. Queensland University of Technology, 2011, pp. 1–16.

[41] A. Choudhary, W. Hendrix, K. Lee, D. Palsetia, and W.-K. Liao, “Social media evolution of the egyptian revolution,” *Communications of the ACM*, vol. 55, no. 5, pp. 74–80, 2012.

[42] J. Cohen, “A coefficient of agreement for nominal scales. educational and psychosocial measurement, 20, 37-46,” 1960.

[43] S. E. Cohen, “Sandy marked a shift for social media use in disasters,” *Emergency Management*, 2013.

[44] K. L. Colley and A. Collier, “An overlooked social media tool? making a case for wikis,” *Public Relations Strategist*, pp. 34–35, 2009.

[45] F. Comunello, *Social Media e comunicazione d’emergenza*, 1st ed. Milano, Italy: Gurini Editore, 6 2014.

[46] M. Conover, J. Ratkiewicz, M. R. Francisco, B. Gonçalves, F. Menczer, and A. Flammini, “Political polarization on twitter.” *ICWSM*, vol. 133, pp. 89–96, 2011.

[47] W. J. Corvey, S. Verma, S. Vieweg, M. Palmer, and J. H. Martin, “Foundations of a multilayer annotation framework for twitter communications during crisis events.” in *LREC*. Citeseer, 2012, p. 2012.

[48] S. Cresci, M. Tesconi, A. Cimino, and F. Dell’Orletta, “A linguistically-driven approach to cross-event damage assessment of natural disasters from social media messages,” in *Proceedings of the 24th International Conference on World Wide Web*. ACM, 2015, pp. 1195–1200.

[49] R. Cross, "The case for integrating crisis response with social media," *Red Cross White Paper*, 2010.

[50] A. Crowe, "Disaster recovery 2.0: The impact of social media on dr processes," in *Facility Fusion 2012*. Ifma, 2012.

[51] A. Culotta, "Towards detecting influenza epidemics by analyzing twitter messages," in *Proceedings of the first workshop on social media analytics*. ACM, 2010, pp. 115–122.

[52] S. Dashti, L. Palen, M. P. Heris, K. M. Anderson, S. Anderson, and S. Anderson, "Supporting disaster reconnaissance with social media data: a design-oriented case study of the 2013 colorado floods," in *Proceedings of the 11th International ISCRAM Conference*, 2014, pp. 18–21.

[53] O. DeMasi, D. Mason, and J. Ma, "Understanding communities via hashtag engagement: A clustering based approach." in *ICWSM*, 2016, pp. 102–111.

[54] W. R. Dombrowsky, "Again and again: Is a disaster what we call a 'disaster'?" *International Journal of Mass Emergencies and Disasters*, vol. 13, no. 3, pp. 241–254, 1995.

[55] T. E. Drabek, "Managing the emergency response," *Public Administration Review*, vol. 45, pp. 85–92, 1985.

[56] ——, *Human system responses to disaster: An inventory of sociological findings*. Springer Science & Business Media, 2012.

[57] T. E. Drabek and D. A. McEntire, "Emergent phenomena and the sociology of disaster: lessons, trends and opportunities from the research literature," *Disaster Prevention and Management: An International Journal*, vol. 12, no. 2, pp. 97–112, 2003.

[58] L. Dwarakanath, "{#} Emergency : Role of Twitter Hashtags during and after a disaster," vol. 3, no. 4, pp. 948–953, 2014.

[59] R. R. Dynes, "Community emergency planning: False assumptions and inappropriate analogies, 1994," *Newark, Delaware: University of Delaware, Disaster Research Center Article*, no. 275, 1994.

[60] ——, *Organized behavior in disaster*. Heath LexingtonBooks, 1970.

[61] P. S. Earle, D. C. Bowden, and M. Guy, "Twitter earthquake detection: earthquake monitoring in a social world," *Annals of Geophysics*, vol. 54, no. 6, 2012.

[62] K. Eismann, O. Posegga, and K. Fischbach, "Collective behaviour, social media, and disasters: A systematic literature review." ECIS, 2016.

[63] M. Eriksson and E.-K. Olsson, "Facebook and twitter in crisis communication: A comparative study of crisis communication professionals and citizens," *Journal of Contingencies and Crisis Management*, vol. 24, no. 4, pp. 198–208, 2016.

[64] F. Faccioli, *Tecnologie digitali per la comunicazione pubblica.* Bonanno, 2013, ch. Tra promozione d'immagine, cultura del servizio e partecipazione civica. Scenari della comunicazione pubblica in Italia.

[65] H. W. Fischer, *Response to disaster: Fact versus fiction & its perpetuation: The sociology of disaster.* University press of America, 1998.

[66] ——, "The sociology of disaster: definitions, research questions, & measurements. continuation of the discussion in a post-september 11 environment," *International Journal of Mass Emergencies and Disasters*, vol. 21, no. 1, pp. 91–108, 2003.

[67] J. D. Fraustino, B. Liu, and Y. Jin, "Social media use during disasters: a review of the knowledge base and gaps." 2012.

[68] L. Freeman, "The development of social network analysis," *A Study in the Sociology of Science*, 2004.

[69] C. E. Fritz, "Contemporary social problems: an introduction to the sociology of deviant behavior and social disorganization," R. K. Merton and R. A. Nisbet, Eds. New York: Brace and World Inc, 1961, ch. Disasters, pp. 651–694.

[70] C. E. Fritz and J. H. Mathewson, *Convergence behavior in disasters: A problem in social control: A special report prepared for the Committee on Disaster Studies.* National Academy of Sciences National Research Council, 1957.

[71] H. Gao, G. Barbier, R. Goolsby, and D. Zeng, "Harnessing the crowdsourcing power of social media for disaster relief," DTIC Document, Tech. Rep., 2011.

[72] F. Giglietto and A. Lovari, "Amministrazioni pubbliche e gestione degli eventi critici attraverso i social media: il caso di# firenzeneve," *Mediascapes journal*, no. 1, pp. 99–118, 2013.

[73] H. Gil de Zúñiga, N. Jung, and S. Valenzuela, "Social media use for news and individuals' social capital, civic engagement and political participation," *Journal of Computer-Mediated Communication*, vol. 17, no. 3, pp. 319–336, 2012.

[74] B. Gleason, "# occupy wall street: Exploring informal learning about a social movement on twitter," *American Behavioral Scientist*, p. 0002764213479372, 2013.

[75] S. González-Bailón, N. Wang, A. Rivero, J. Borge-Holthoefer, and Y. Moreno, "Assessing the bias in communication networks sampled from twitter," *arXiv preprint arXiv:1212.1684*, 2012.

[76] M. F. Goodchild, "Citizens as sensors: the world of volunteered geography," *GeoJournal*, vol. 69, no. 4, pp. 211–221, 2007.

[77] M. F. Goodchild and J. A. Glennon, "Crowdsourcing geographic information for disaster response: a research frontier," *International Journal of Digital Earth*, vol. 3, no. 3, pp. 231–241, 2010.

[78] J. Gottfried and E. Shearer, "News use across social media platforms 2016," *Pew Research Center*, 2016.

[79] V. Grasso and A. Crisci, "Codified hashtags for weather warning on twitter: an italian case study," *PLOS Currents Disasters*, 2016.

[80] V. Grasso, F. Zabini, A. Crisci, and M. Perna, "Mappare la partecipazione. l'impronta geografica degli utenti digitali del lamma, servizio meteo della toscana," in *Atti 18a Conferenza Nazionale ASITA*, vol. 1416. ASITA, 2014, pp. 635–642.

[81] V. S. M. W. Group and D. F. R. Group, "Lessons learned: Social media and hurricane sandy," U.S. Department of Homelenad and Security, Washington D.C., United States, Tech. Rep., June 2013. [Online]. Available: https://www.dhs.gov/sites/default/files/publications/Lessons%20Learned%20Social%20Media%20and%20Hurricane%20Sandy.pdf

[82] A. Gruzd, B. Wellman, and Y. Takhteyev, "Imagining twitter as an imagined community," *American Behavioral Scientist*, vol. 55, no. 10, pp. 1294–1318, 2011.

[83] C. Haythornthwaite, "Social networks and internet connectivity effects," *Information, Community & Society*, vol. 8, no. 2, pp. 125–147, 2005.

[84] I. Helsloot and A. Ruitenberg, "Citizen response to disasters: a survey of literature and some practical implications," *Journal of Contingencies and Crisis Management*, vol. 12, no. 3, pp. 98–111, 2004.

[85] K. Hewitt, "Interpretations of calamity from the viewpoint of human ecology," 1983.

[86] ——, "Excluded perspectives in the social construction of disaster," *What is a disaster*, pp. 75–91, 1998.

[87] J. Horrigan, *Relief donations after Hurricanes Katrina and Rita and use of the Internet to get disaster news.* Pew Internet & American Life Project, 2005.

[88] J. B. Houston, J. Hawthorne, M. F. Perreault, E. H. Park, M. Goldstein Hode, M. R. Halliwell, S. E. Turner McGowen, R. Davis, S. Vaid, J. A. McElderry *et al.*, "Social media and disasters: a functional framework for social media use in disaster planning, response, and research," *Disasters*, vol. 39, no. 1, pp. 1–22, 2015.

[89] C.-M. Huang, E. Chan, and A. A. Hyder, "Web 2.0 and internet social networking: A new tool for disaster management?-lessons from taiwan," *BMC medical informatics and decision making*, vol. 10, no. 1, p. 57, 2010.

[90] A. L. Hughes and L. Palen, "The evolving role of the public information officer: An examination of social media in emergency management," *Journal of Homeland Security and Emergency Management*, vol. 9, no. 1, pp. 1–20, 2012.

[91] A. L. Hughes, L. Palen, J. Sutton, S. B. Liu, and S. Vieweg, "Site-seeing in disaster: An examination of on-line social convergence," in *Proceedings of the 5th International ISCRAM Conference.* Citeseer, 2008.

[92] A. L. Hughes, L. A. St Denis, L. Palen, and K. M. Anderson, "Online public communications by police & fire services during the 2012 hurricane sandy," in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems.* ACM, 2014, pp. 1505–1514.

[93] A. L. Hughes and L. Palen, "Twitter adoption and use in mass convergence and emergency events," *International Journal of Emergency Management*, vol. 6, no. 3-4, pp. 248–260, 2009.

[94] N. Ireson, "Local community situational awareness during an emergency," in *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies.* IEEE, 2009, pp. 49–54.

[95] U. ISDR, "Unisdr terminology on disaster risk reduction," *Geneva, Switzerland, May*, 2009.

[96] A. M. Italiana, "Bollettino mensile novembre 2014," Aeronautica Militare Italiana, Roma, Tech. Rep., 11 2014.

[97] M. Jacomy, S. Heymann, T. Venturini, and M. Bastian, "Forceatlas2, a continuous graph layout algorithm for handy network visualization," *Medialab center of research*, vol. 560, 2011.

[98] P. T. Jaeger, B. Shneiderman, K. R. Fleischmann, J. Preece, Y. Qu, and P. F. Wu, "Community response grids: E-government, social networks, and effective emergency management," *Telecommunications Policy*, vol. 31, no. 10, pp. 592–604, 2007.

[99] Y. Jin, B. F. Liu, and L. L. Austin, "Examining the role of social media in effective crisis management: The effects of crisis origin, information

form, and source on publics' crisis responses," *Communication Research*, p. 0093650211423918, 2011.

[100] D. Johnson, A. Zagorecki, J. M. Gelman, and L. K. Comfort, "Improved situational awareness in emergency management through automated data analysis and modeling," *Journal of Homeland Security and Emergency Management*, vol. 8, no. 1, 2011.

[101] T. Johnson and B. Kaye, "Choosing is believing? how web gratifications and reliance affect internet credibility among politically interested users," *Atlantic Journal of Communication*, vol. 18, no. 1, pp. 1–21, 2010.

[102] J. Jones, "Switching in twitter's hashtagged exchanges," *Journal of Business and Technical Communication*, vol. 28, no. 1, pp. 83–108, 2014.

[103] D. Jurgens, T. Finethy, J. McCorriston, Y. T. Xu, and D. Ruths, "Geolocation prediction in twitter using social networks: A critical analysis and review of current practice," in *Proceedings of the 9th International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2015.

[104] L. Kaewkitipong, C. Chen, and P. Ractham, "Lessons learned from the use of social media in combating a crisis: A case study of 2011 thailand flooding disaster," 2012.

[105] E. Katz and D. Foulkes, "On the use of the mass media as "escape": Clarification of a concept," *Public opinion quarterly*, vol. 26, no. 3, pp. 377–388, 1962.

[106] A. L. Kavanaugh, E. A. Fox, S. D. Sheetz, S. Yang, L. T. Li, D. J. Shoemaker, A. Natsev, and L. Xie, "Social media use by government: From the routine to the critical," *Government Information Quarterly*, vol. 29, no. 4, pp. 480–491, 2012.

[107] P. Kennedy, E. Ressler, H. Rodriguez, E. L. Quarantelli, R. Dynes *et al.*, *Handbook of disaster research.* Springer Science & Business Media, 2009.

[108] L. M. Killian, "Some accomplishments and some needs in disaster study," *Journal of social Issues*, vol. 10, no. 3, pp. 66–72, 1954.

[109] K. Kodrich and M. Laituri, "Making a connection: Social media's key role in the haiti earthquake," *Journal of Communication and Computer*, vol. 8, no. 8, pp. 624–627, 2011.

[110] A. Kongthon, C. Haruechaiyasak, J. Pailai, and S. Kongyoung, "The role of twitter during a natural disaster: Case study of 2011 thai flood," in *2012 Proceedings of PICMET'12: Technology Management for Emerging Technologies.* IEEE, 2012, pp. 2227–2232.

[111] S. M. Kywe, T.-A. Hoang, E.-P. Lim, and F. Zhu, "On recommending hashtags in twitter networks," in *International Conference on Social Informatics.* Springer, 2012, pp. 337–350.

[112] LaMMA, "Report evento 10-11 novembre 2014," Consorzio LaMMA, Regione Toscana - CNR, Sesto Fiorentino, Firenze, Tech. Rep., 11 2014.

[113] ——, "Report evento 4-5 novembre 2014," Consorzio LaMMA, Regione Toscana - CNR, Sesto Fiorentino, Firenze, Tech. Rep., 11 2014.

[114] W. R. Lariscy, S. F. Tinkham, and K. D. Sweetser, "Kids these days: Examining differences in political uses and gratifications, internet political participation, political information efficacy, and cynicism on the basis of age," *American Behavioral Scientist*, vol. 55, no. 6, pp. 749–764, 2011.

[115] M. Latonero and I. Shklovski, "Emergency management, twitter, and social media evangelism," *Latonero, M. & Shklovski, I.(2011). Emergency management, Twitter, & Social Media Evangelism. International Journal of Information Systems for Crisis Response and Management*, vol. 3, no. 4, pp. 67–86, 2011.

[116] A. Lev-On, "Communication, community, crisis: Mapping uses and gratifications in the contemporary media environment," *new media & society*, vol. 14, no. 1, pp. 98–116, 2012.

[117] Y.-R. Lin, D. Margolin, B. Keegan, A. Baronchelli, and D. Lazer, "# bigbirds never die: understanding social dynamics of emergent hashtag," *arXiv preprint arXiv:1303.7144*, 2013.

[118] B. F. Liu, Y. Jin, and L. L. Austin, "The tendency to tell: Understanding publics' communicative responses to crisis information form and source," *Journal of Public Relations Research*, vol. 25, no. 1, pp. 51–67, 2013.

[119] B. Longueville, G. Luraschi, P. Smits, S. Peedell, and T. Groeve, "Citizens as sensors for natural hazards: A vgi integration workflow," *Geomatica*, vol. 64, pp. 41–59, 2010.

[120] A. Lovari, *Networked citizens: comunicazione pubblica e amministrazioni digitali.* F. Angeli, 2013.

[121] ——, "Comunicatori pubblici dietro la timeline. la sfida dei social media tra nuove competenze professionali e vecchie resistenze al cambiamento," *Problemi dell'informazione*, vol. 41, no. 1, pp. 61–86, 2016.

[122] A. Majchrzak and P. H. More, "Emergency! web 2.0 to the rescue!" *Communications of the ACM*, vol. 54, no. 4, pp. 125–132, 2011.

[123] L. Materassi, "Intermediari e pratiche di disintermediazione. nuove tecnologie e social media nei comuni toscani," *Problemi dell'informazione*, vol. 41, no. 1, pp. 87–112, 2016.

[124] L. Materassi and L. Solito, "Communication's routes in public administration: institutional arrangements in tuscany," *Problemi dell'informazione*, vol. 40, no. 2, pp. 303–334, 2015.

[125] P. Meier, "Human computation for disaster response," in *Handbook of human computation*. Springer, 2013, pp. 95–104.

[126] ——, *Digital humanitarians: how big data is changing the face of humanitarian response*. Crc Press, 2015.

[127] N. Merrifield and M. Panechar, "Uncertainty reduction strategies via twitter: The 2011 wildfire threat to los alamos national laboratory," in *Proceedings from AEJMC Annual Conference*, 2012.

[128] M. M. Miglietta and R. Rotunno, "An ef3 multivortex tornado over the ionian region: Is it time for a dedicated warning system over italy?" *Bulletin of the American Meteorological Society*, vol. 97, no. 3, pp. 337–344, 2016.

[129] D. Mileti, *Disasters by design: A reassessment of natural hazards in the United States*. Joseph Henry Press, 1999.

[130] D. S. Mileti, T. E. Drabek, and J. E. Haas, *Human systems in extreme environments: A sociological perspective*. Institute of Behavioral Science, University of Colorado, 1975, vol. 21.

[131] D. S. Mileti and J. H. Sorensen, "Communication of emergency public warnings," *Landslides*, vol. 1, no. 6, pp. 52–70, 1990.

[132] A. Mills, R. Chen, J. Lee, and H. Raghav Rao, "Web 2.0 emergency applications: How useful can twitter be for emergency response?" *Journal of Information Privacy and Security*, vol. 5, no. 3, pp. 3–26, 2009.

[133] H. E. Moore, *Tornadoes Over Texas*. University of Texas Press, 1958.

[134] J. L. Moreno, "Who shall survive? foundations of sociometry, group psychotherapy and socio-drama," 1953.

[135] J. Moreno, "Who shall survive? foundations of sociometry, group psychotherapy, and sociodrama. beacon, ny: Beacon house," 1934.

[136] R. Munro, L. Gunasekara, S. Nevins, L. Polepeddi, and E. Rosen, "Tracking epidemics with natural language processing and crowdsourcing," in *2012 AAAI Spring Symposium Series*, 2012.

[137] P. Nesi, G. Pantaleo, and G. Sanesi, "A hadoop based platform for natural language processing of web pages and documents," *Journal of Visual Languages & Computing*, vol. 31, pp. 130–138, 2015.

[138] U. OCHA, "Hashtag standards for emergencies," Technical Report. United Nations Office for the Coordination of Humanitarian Affairs, Tech. Rep., 2014.

[139] A. Olteanu, S. Vieweg, and C. Castillo, "What to expect when the unexpected happens: Social media communications across crises," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2015, pp. 994–1009.

[140] L. Palen, K. M. Anderson, G. Mark, J. Martin, D. Sicker, M. Palmer, and D. Grunwald, "A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters," in *Proceedings of the 2010 ACM-BCS visions of computer science conference.* British Computer Society, 2010, p. 8.

[141] L. Palen and S. B. Liu, "Citizen communications in crisis: anticipating a future of ict-supported public participation," in *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 2007, pp. 727–736.

[142] L. Palen, S. Vieweg, S. B. Liu, and A. L. Hughes, "Crisis in a networked world features of computer-mediated communication in the april 16, 2007, virginia tech event," *Social Science Computer Review*, vol. 27, no. 4, pp. 467–480, 2009.

[143] L. Parisi, F. Comunello, and A. Amico, *Social media e comunicazione di emergenza.* Milano: Guerini Editore, 2014, ch. #allertameteoSAR: analisi di un hashtag di servizio tra dinamiche di influenza e nuove forme di engagement.

[144] S.-Y. Perng, M. Buscher, R. Halvorsrud, L. Wood, M. Stiso, L. Ramirez, and A. Al-Akkad, "Peripheral response: Microblogging during the 22/7/2011 norway attacks," 2012.

[145] R. W. Perry, "What is a disaster?" *Handbook of disaster research*, pp. 1–15, 2007.

[146] R. W. Perry and M. K. Lindell, "Understanding citizen response to disasters with implications for terrorism," *Journal of Contingencies and Crisis Management*, vol. 11, no. 2, pp. 49–60, 2003.

[147] M. Porta, *A dictionary of epidemiology.* Oxford university press, 2014.

[148] C. Prell, *Social network analysis: History, theory and methodology.* Sage, 2012.

[149] C. H. Procopio and S. T. Procopio, "Do you know what it means to miss new orleans? internet communication, geographic community, and social capital in crisis," *Journal of Applied Communication Research*, vol. 35, no. 1, pp. 67–87, 2007.

[150] E. L. Quarantelli, "Organizational behavior in disasters and implications for disaster planning. volume 1, number 2," DTIC Document, Tech. Rep., 1986.

[151] ——, "Conceptualizing disasters from a sociological perspective," 1989.

[152] E. L. Quarantelli, P. Lagadec, and A. Boin, "A heuristic approach to future disasters and crises: new, old, and in-between types," *Handbook of disaster research*, pp. 16–41, 2007.

[153] E. L. Quarantelli, *Organization under stress.* Disaster Research Center, the Ohio State University, 1966.

[154] ——, "Disaster studies: An analysis of the social historical factors affecting the development of research in the area," 1987.

[155] ——, "Emergencies, disasters and catastrophes are different phenomena," 2000.

[156] ——, *What is a disaster?: a dozen perspectives on the question.* Routledge, 2005.

[157] 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, no. 4, pp. 919–931, 2013.

[158] J. M. Santos, B. E. Aguirre *et al.*, "Communicating risk and uncertainty: science, technology, and disasters at the crossroads," in *Handbook of disaster research.* Springer, 2007, pp. 476–488.

[159] A. Sarcevic, L. Palen, J. White, K. Starbird, M. Bagdouri, and K. Anderson, "Beacons of hope in decentralized coordination: Learning from on-the-ground medical twitterers during the 2010 haiti earthquake," in *Proceedings of the ACM 2012 conference on computer supported cooperative work.* ACM, 2012, pp. 47–56.

[160] J. Scanlon, "Unwelcome irritant or useful ally? the mass media in emergencies," *Handbook of Disaster research*, pp. 413–429, 2007.

[161] J. Scanlon, I. Helsloot, and J. Groenendaal, "Putting it all together: Integrating ordinary people into emergency response." *International Journal of Mass Emergencies & Disasters*, vol. 32, no. 1, 2014.

[162] J. Scott and P. J. Carrington, *The SAGE handbook of social network analysis.* SAGE publications, 2011.

[163] G. Shao, "Understanding the appeal of user-generated media: a uses and gratification perspective," *Internet Research*, vol. 19, no. 1, pp. 7–25, 2009.

[164] C. Shirky, *Here comes everybody: The power of organizing without organizations.* Penguin, 2008.

[165] I. Shklovski, L. Palen, and J. Sutton, "Finding community through information and communication technology in disaster response," in *Proceedings of the 2008 ACM conference on Computer supported cooperative work.* ACM, 2008, pp. 127–136.

[166] F. Silvestro, N. Rebora, F. Giannoni, A. Cavallo, and L. Ferraris, "The flash flood of the bisagno creek on 9th october 2014: An unfortunate combination of spatial and temporal scales," *Journal of Hydrology*, 2015.

[167] B. G. Smith, "Socially distributing public relations: Twitter, haiti, and interactivity in social media," *Public Relations Review*, vol. 36, no. 4, pp. 329–335, 2010.

[168] M. H. Smith, "American religious organizations in disaster: A study of congregational response to disaster," *Mass Emergencies*, vol. 3, no. 2-3, pp. 133–142, 1978.

[169] L. Solito, "Tra cambiamenti visibili e immobilismi opachi: la comunicazione pubblica in italia," *Sociologia della comunicazione*, 2014.

[170] C. R. Soriano, M. D. G. Roldan, C. Cheng, and N. Oco, "Social media and civic engagement during calamities: the case of twitter use during typhoon yolanda," *Philippine Political Science Journal*, vol. 37, no. 1, pp. 6–25, 2016.

[171] E. S. Spiro, S. Fitzhugh, J. Sutton, N. Pierski, M. Greczek, and C. T. Butts, "Rumoring during extreme events: A case study of deepwater horizon 2010," in *Proceedings of the 4th Annual ACM Web Science Conference*. ACM, 2012, pp. 275–283.

[172] R. A. Stallings, "The structural patterns of four types of organizations in disaster," *Disasters: Theory and research*, pp. 87–103, 1978.

[173] ——, "Disaster and the theory of social order," *What is a Disaster*, pp. 127–145, 1998.

[174] R. A. Stallings and E. L. Quarantelli, "Emergent citizen groups and emergency management," *Public administration review*, vol. 45, pp. 93–100, 1985.

[175] K. Starbird and L. Palen, *Pass it on?: Retweeting in mass emergency*. International Community on Information Systems for Crisis Response and Management, 2010.

[176] ——, "Voluntweeters: Self-organizing by digital volunteers in times of crisis," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 1071–1080.

[177] K. Starbird, L. Palen, A. L. Hughes, and S. Vieweg, "Chatter on the red: what hazards threat reveals about the social life of microblogged information," in *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. ACM, 2010, pp. 241–250.

[178] K. Starbird and J. Stamberger, "Tweak the tweet: Leveraging microblogging proliferation with a prescriptive syntax to support citizen reporting," 2010.

[179] K. K. Stephens and P. C. Malone, "If the organizations won't give us information: The use of multiple new media for crisis technical translation and dialogue," *Journal of Public Relations Research*, vol. 21, no. 2, pp. 229–239, 2009.

[180] S. Stieglitz, L. Dang-Xuan, A. Bruns, and C. Neuberger, "Social media analytics," *Wirtschaftsinformatik*, vol. 56, no. 2, pp. 101–109, 2014.

[181] B. Suh, L. Hong, P. Pirolli, and E. H. Chi, "Want to be retweeted? large scale analytics on factors impacting retweet in twitter network," in *Social computing (socialcom), 2010 ieee second international conference on*. IEEE, 2010, pp. 177–184.

[182] J. Sutton, "When online is off: Public communications following the february 2011 christchurch, nz, earthquake," in *Proceedings of the 9th International Conference on Information Systems for Crisis Response and Management*, 2012.

[183] J. Sutton, L. Palen, and I. Shklovski, "Backchannels on the front lines: Emergent uses of social media in the 2007 southern california wildfires," in *Proceedings of the 5th International ISCRAM Conference*. Washington, DC, 2008, pp. 624–632.

[184] J. Sutton, E. S. Spiro, B. Johnson, S. Fitzhugh, B. Gibson, and C. T. Butts, "Warning tweets: Serial transmission of messages during the warning phase of a disaster event," *Information, Communication & Society*, vol. 17, no. 6, pp. 765–787, 2014.

[185] K. D. Sweetser and E. Metzgar, "Communicating during crisis: Use of blogs as a relationship management tool," *Public Relations Review*, vol. 33, no. 3, pp. 340–342, 2007.

[186] M. Taylor and M. L. Kent, "Taxonomy of mediated crisis responses," *Public Relations Review*, vol. 33, no. 2, pp. 140–146, 2007.

[187] M. Taylor, G. Wells, G. Howell, B. Raphael *et al.*, "The role of social media as psychological first aid as a support to community resilience building," *Australian Journal of Emergency Management, The*, vol. 27, no. 1, p. 20, 2012.

[188] I. Temnikova, S. Vieweg, and C. Castillo, "The case for readability of crisis communications in social media," in *Proceedings of the 24th International Conference on World Wide Web*. ACM, 2015, pp. 1245–1250.

[189] C. Torrey, M. Burke, M. Lee, A. Dey, S. Fussell, and S. Kiesler, "Connected giving: Ordinary people coordinating disaster relief on the internet," in *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*. IEEE, 2007, pp. 179a–179a.

[190] J. Travers and S. Milgram, "The small world problem," *Phychology Today*, vol. 1, pp. 61–67, 1967.

[191] S. R. Veil, T. Buehner, and M. J. Palenchar, "A work-in-process literature review: Incorporating social media in risk and crisis communication," *Journal of contingencies and crisis management*, vol. 19, no. 2, pp. 110–122, 2011.

[192] S. Verma, S. Vieweg, W. Corvey, L. Palen, J. H. Martin, M. Palmer, A. Schram, and K. M. Anderson, "Nlp to the rescue?: Extracting" situational awareness" tweets during mass emergency," in *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[193] S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen, "Microblogging during two natural hazards events: what twitter may contribute to situational awareness," in *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2010, pp. 1079–1088.

[194] G. Visconti and F. S. Marzano, "An independent overview of the national weather service in italy," *Bulletin of the American Meteorological Society*, vol. 89, no. 9, p. 1279, 2008.

[195] A. F. Wallace, *Human behavior in extreme situations: a survey of the literature and suggestions for further research*. National Academy of Sciences, 1956.

[196] D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" networks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[197] D. E. Wenger, "The study of volunteer and emergent organizational response: Approaches and issues for future research," in *En*, 1989, pp. 350–365.

[198] ——, *Emergent and volunteer behavior during disaster: research findings and planning implications*. Texas A & M University. Hazard Reduction & Recovery Center, 1992.

[199] E. Wenger and B. Wenger-Trayner, "Introduction to communities of practice. a brief overview of the concept and its uses," *Online: http://wenger-trayner.com/introduction-to-communitiesof-practice*, 2015.

[200] A. Whiting and D. Williams, "Why people use social media: a uses and gratifications approach," *Qualitative Market Research: An International Journal*, vol. 16, no. 4, pp. 362–369, 2013.

[201] J. Whittaker, B. McLennan, and J. Handmer, "A review of informal volunteerism in emergencies and disasters: Definition, opportunities and challenges," *International journal of disaster risk reduction*, vol. 13, pp. 358–368, 2015.

[202] Wikipedia, "Twitter — wikipedia, the free encyclopedia," 2017, [Online; accessed 9-May-2017]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Twitter&oldid=779444901

[203] H. B. Williams, "Human factors in warning and response systems," *The threat of impending disaster*, pp. 96–109, 1964.

[204] C. Wukich and A. Steinberg, "Nonprofit and public sector participation in self-organizing information networks: Twitter hashtag and trending topic

use during disasters," *Risk, Hazards & Crisis in Public Policy*, vol. 4, no. 2, pp. 83–109, 2013.

[205] L. Yang, T. Sun, M. Zhang, and Q. Mei, "We know what@ you# tag: does the dual role affect hashtag adoption?" in *Proceedings of the 21st international conference on World Wide Web*. ACM, 2012, pp. 261–270.

[206] S. Yardi and D. Boyd, "Dynamic debates: An analysis of group polarization over time on twitter," *Bulletin of Science, Technology & Society*, vol. 30, no. 5, pp. 316–327, 2010.

[207] D. Yates and S. Paquette, "Emergency knowledge management and social media technologies: A case study of the 2010 haitian earthquake," *International journal of information management*, vol. 31, no. 1, pp. 6–13, 2011.

[208] M. Zappavigna, "Searchable talk: the linguistic functions of hashtags," *Social Semiotics*, vol. 25, no. 3, pp. 274–291, 2015.

[209] L. Zheng and T. Zheng, "Innovation through social media in the public sector: Information and interactions," *Government Information Quarterly*, vol. 31, pp. S106–S117, 2014.

[210] M. Zook, M. Graham, T. Shelton, and S. Gorman, "Volunteered geographic information and crowdsourcing disaster relief: a case study of the haitian earthquake," *World Medical & Health Policy*, vol. 2, no. 2, pp. 7–33, 2010.