

# Local Community Situational Awareness During An Emergency

Neil Ireson<sup>1</sup>

<sup>1</sup>Neil Ireson, Department of Computer Science, Sheffield University, Sheffield, UK.  
e-mail : n.ireson@dcs.shef.ac.uk

*With the recent exponential growth in the use of public forums for communication and sharing of information, Emergency Response organisations are realising that there is a potential to exploit these forums to extract information to improve the overall awareness of the events that occur during an emergency. This paper describes an approach using information extraction, topic and event identification and its application in a case study. The results indicate that despite the ill-formed, variable quality and conversational nature of forum posts a degree of coherent event information can be acquired.*

*Emergency Response, Information Extraction, Community Information.*

## I. INTRODUCTION

There are many different types of emergencies: ones brought about by forces of nature such as avalanches, floods, droughts, earthquakes or man-made emergencies, e.g. train and plane crashes, pollution and terrorist attacks. These emergencies can vary in terms of scale, both in severity and affected location. In small scale emergencies, only a few people and organisations may be involved, typically local authorities such as the city council, police and fire department. During large-scale emergencies several hundred organisations can be involved, and thousands or even millions of individuals; as was the situation after the Tsunami in the Indian Ocean 2004.

The Emergency Response (ER) team's decision-making process primarily involves the allocation and coordination of resources, but also maintaining effective communication between: the organisations involved, the decision/command chain and the affected individuals. During an incident, the ER team receives information from multifarious sources (e.g. the fire/police/ambulance services, other local authority bodies, government bodies, broadcast services, affected individuals, etc.). The management of this mass of information is crucial in aiding the decision-making process, ensuring, as far as possible, that the responders have full *situational awareness* to make informed decisions.

Situational awareness can be defined as comprehension of what is happening around you and the ability to retrieve, understand and reuse the available information to enhance your response to the situation. It is of particular relevance when dealing with large amount of information and when the consequences of actions may be dramatic. In ER, situational awareness can be related to the possibility of having accurate, complete and real-time information about an incident. The seriousness of an incident is likely to increase as its scale and geographic distribution increases; in such incidents situational awareness becomes increasingly challenging and important, as limited resources and communication

channels become stretched.

The advent of the Internet as a new media channel offers individuals the possibility of instantly sharing and broadcasting personal content (text, images, videos). Although people tend to communicate information related to their individual experience (mostly with family, friends or known colleagues), where this communication occurs during an emergency the citizens in the community provide a potentially valuable source of collective information, via their conversational interactions. This information can be extracted from each individual's contribution and combined to generate an emergent picture of the events in an emergency.

The work described in this paper forms part of the WeKnowIt<sup>1</sup> European Research Project, which aims to develop methodologies and tools enabling the exploitation of this collective intelligence to benefit the organisations and communities involved in the emergency.

The following sections describes the nature of this community information, including the potentially issues arising from its use, and the application of Information Extraction and Event Detection techniques to explicit extract the information implicitly provided by the community.

## II. COMMUNITY INFORMATION

The information requirements of ER organisations can be very broadly summarised by the classical facets of information gathering; *who* is involved in the incident, *what* type of incident is it, *where* is the incident taking place and *when* did the incident start, and end. The recent exponential growth in user-generated content sites ([2][6]) is providing text, images and videos which potentially offer useful information to meet some of these requirements. Rather than relying on the limited resources of ER organisations to gather information, social network sites provide access to a mass of individuals who are directly involved in the incident. In addition, with the advent of mobile interfaces to social networks, users can upload information directly from the site of the incident, thus providing real-time critical information about the event, and the possibility of having a clearer geographic visualisation of the extent of the emergency.

In fact, for a number of recent earthquakes it has been claimed that Twitter<sup>2</sup>, a micro-blogging service, provided the first notification and pictures related to seismic events before the national broadcast services or even professional ER organisations<sup>3</sup>. In addition, after the mass shooting at

<sup>1</sup> <http://www.weknowit.eu/>

<sup>2</sup> <http://twitter.com/>

<sup>3</sup> [http://www.bbc.co.uk/blogs/technology/2008/05/twitter\\_and\\_the\\_china\\_earthqua.html](http://www.bbc.co.uk/blogs/technology/2008/05/twitter_and_the_china_earthqua.html)

Virginia Tech in April 2007 the online social network had correctly identified the 32 victims before the university released their names [13].

In practice, during an emergency incident, individuals involved in the incident use the communities they are already engaged in as a means to express their concerns and provide or request information. As part of the user requirements gathering phase of the EU project WeKnowIt [12], a questionnaire distributed to local citizens about the Sheffield floods in 2007 showed that 31% of respondents used social networks and public forums as a mean of communication; when asked to only consider communication with people outside their family this figure rises to 40%. This indicates that social networks are becoming a common resource for citizens during an emergency and thus that a significant proportion of the population is likely to use social sites to both provide and receive information.

Whilst ER organisations engage in knowledge management processes aiming to generate and distribute accurate information, user-generated content may not adhere to the same criteria. Users may post information that is speculative rather than definitive, or simply incorrect or misleading. Where incorrect information is transferred from a social network setting and used or reproduced by an organisation the effect of any misinformation may be more serious, for example leading to inappropriate resource allocation or loss of the organisation's reputation. In addition to the possibility of incorrect information, the fact that nature of information is subjective means individuals and ER organisations will have different perspectives and place different degrees of importance upon a given incident. To an extent, pictorial information can be seen to be more objective and reliable; by avoiding editorial bias and allowing ER personnel to more directly assess the degree of damage or danger. However, care is still required, as shown by the case where Sky News published user submitted pictures of the recent UK storms that included ones from the New Orleans floods and stills from the film "The Day After Tomorrow"<sup>4</sup>.

If organisations are going to exploit user-generated content from social networks, they need means to assess the quality of that information. One of the key ways of assessment of information is independent validation; this might come from other users or external sources. As given users provide quality information, the trust in those users grows, and this trust can be propagated to their associates. For organisations, which may have multiple people or even automatic systems interacting with the social network, this network of trust will have to be automatically determined rather than inferred via subjective assessment built up from historic interactions. Although individuals have their own agenda and interests, the summation of their contributions can lead to a collective intelligence, which provides information not directly related to the intention of the individuals in the community.

The use of social networks, and Web 2.0 platforms, will radically affect the way the ER information is distributed: from a one-way communication paradigm (where an organisation gathers information from internal or trusted sources

and communicates this to the citizens) towards two-way communication in which the citizens become active members of the ER team with respect to online information gathering and communication. The use of user-generated content poses issues regarding the quality and the trustworthiness of information, as the ER organisations no longer have full control over the information. Therefore the concept of control may well need to be revised into the "management and cooperation" in the gathering and communication of information.

### III. INFORMATION EXTRACTION

The process of extracting the desired information from open discussion forums, according to some predefined purpose, poses several issues for classical information extraction. The techniques employed have largely been developed to identify (linguistic) patterns from large, static, well-edited texts. However the nature of the input to forums, blogs, and other community web sites means the text is noisy, error-ridden, ill-formed and conversational in nature. Whilst posts do have a chronological order, they related to multifarious interleaved conversations between many authors.

The extraction of the information woven into these discussions requires an approach which can handle artefacts such as spelling errors, colloquialisms, (forum) sub-language expressions (lol, omg, smiley's:), etc.

The process of information extraction is divided into five separate sequential tasks:

1. The identification of "topical" groups of posts.
2. Extraction of structured explicit information: post time, user, links (quotation) to other posts.
3. Extraction of unstructured implicit (geographic and temporal) information
4. Identification of sub-topic (event) types
5. Identification of events

#### *Identification of Topical Posts*

Given that we know, a priori, the general topic of interest (i.e. emergency situations) the identification of threads of interest involves looking for keywords or phrases that relate to emergencies: flood, explosion, fire, gas, vehicle accident, etc. Threads that contain these terms are highlighted for further processing.

#### *Structured Information Extraction*

When a post is input into a forum certain structured information (i.e. metadata) is automatically added to it. This includes the time of the posting, the user identification (although open forums do sometimes allow unregistered users to post anonymously) and identification of posts which are explicitly quoted by other posts. This information provides a certain amount of context. The structured representation of this information means it can be extracted using regular pattern matching of the text or/and HTML.

---

<sup>4</sup> <http://socialmediatrader.com/when-user-generated-content-attacks>

## Unstructured Information Extraction

Extracting information from free-text is primarily concerned with named-entity recognition, i.e. the identification of text fragments relating to specific concepts. For the domain of ER temporal and geospatial (location) entities are of primary interest. The approach adopted uses gazetteers of temporal and geospatial terms, these are language and location specific, and combine gazetteer lookup and heuristically derived pattern matching to identify entities.

Temporal entities in the text can either relate to an absolute point in time, i.e. a day, date, month and year, or a relative time expressions, i.e. yesterday, this morning, an hour ago, these are normalised to an absolute time from the time of the post.

For geospatial entities a gazetteer of local place names is used, covering towns and cities, districts and boroughs, street names. Other geographical features can also be referred to, such as parks, pubs, playing fields, and rivers, the latter of these being obviously pertinent to a flooding scenario. The list of place names is compared with the list of common terms, and any co-occurrences are highlighted as potentially ambiguous, thus requiring other confirmatory evidence before being deemed as a geospatial reference.

Once a geographic location is identified in the text there is also the need to geo-code the entity, i.e. associate the location to a unique point (or area) on earth. For the low granularity locations (e.g. cities, districts) there is little ambiguity, however at the level of individual streets there are often a number of possible locations. These can either be different streets, which possess the same name, or one (long) street, which passes through multiple districts. To resolve the ambiguity the context of the entity is used. The disambiguation process constructs a probability distribution of the potential locations based on the other non-ambiguous locations in its context. The probability is increased if the contextual locations do not conflict with the potential location, i.e. they are the same location (or if both locations are streets that they are within some distance threshold) or the potential location is within the non-ambiguous location, if it refers to a wider area. If a single location receives the highest probability, given the context, it is selected as the actual location. The process can be applied more strictly by requiring that the selected most probably location does not conflict with any location in its context.

The disambiguation process uses increasingly distant context: initially considering the other locations in the same post, then locations in directly quoted posts, followed by locations referred to by the same user in other posts and finally locations in preceding posts in the same thread.

### Sub-topic Identification

In order to identify sub-topics the original post text is pre-processed to remove common words, which are deemed as uninformative. In addition, text relating to the geospatial and temporal information, discovered in the processes described above, is removed. The remaining text is compared against text from a collection of general posts from the same forum. The comparison uses a standard natural language processing techniques, i.e. the removal of uninformative

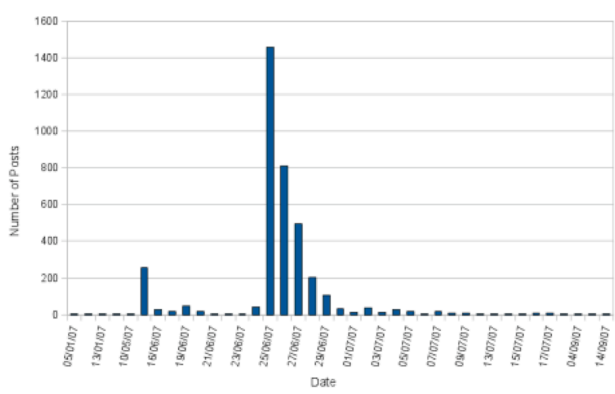


Fig.1 Temporal distribution of flood related posts

terms (stopwords), the normalisation of terms to their stem (applying the Porter Stemming algorithm [8] to remove inflected variation) and the use of a (term frequency / inverse document frequency) measure to determine the important terms, i.e. ones that are more likely to be identified with the topical posts. The terms occurring with a frequency above a given threshold are then clustered together according to their similarity, where the similarity measure is the mean co-occurrence between terms of each cluster (known as average linkage clustering).

A hierarchical clustering approach is then applied, where initially each term is placed in a separate cluster and then the most similar clusters are combined to form a new cluster. The process is iterated until a similarity threshold is reached. Each cluster then provides a list of terms (and related posts) that define the sub-topics.

### Event Identification

The main aim of the information extraction process is to identify groups of posts that refer to a specific (sub-) event. The identification of events is performed using a similar approach to sub-topic identification, i.e. the application of hierarchical clustering. In this case the similarity measure is a function of the geospatial information (proximity to locations referred to in the text) and topic information (the co-occurrence of topic keywords).

## IV. CASE STUDY

In order to explore the extent to which it is possible to build up a coherent picture of an emergency from public forum postings, a case study based on the flooding that took place in Sheffield, UK around June 25<sup>th</sup> 2007, is used. The data in the study has been collected from a local public forum<sup>5</sup>. The “positive” posts, i.e. ones on topic, were determined simply by selecting threads in which the term flood was referred to over 10 times. This resulted in 3621, potentially, flood related posts, each containing on average only 115 characters. The temporal distribution of the flood posts is shown in Figure 1. The main day of the flood, the 25<sup>th</sup>, has almost 1500 flood posts, the average number of posts which the forum usually receives: this is a further confir-

<sup>5</sup> <http://www.sheffieldforum.co.uk/>

mation that public forums are increasingly used as a medium for communication during such an emergency. Note that the slight peak seen a week before relates to a weekend when there was actual more rainfall than on the day of the flooding.

Figure 2 below shows the frequency of post of the 845 users who posted to the flood threads. On average users posted 4.3 posts, and over 50% of posts are from the 10% of users who post 9 times or more. This means that it is the potential to build a profile for a number of users, which provides a context that can be used facilitate the interpretation of information from one of their individual posts, however the majority of users post too infrequently. It is worth noting that these figures only relate to the flood posts; if it were possible to constantly monitor users then more data would be available to construct the profiles.

There are 10 users who have posted over 40 times, with one user posting 198 times (5.4% of the total number of posts). In examining the frequent posting users it can be seen that these individuals tend to provide a key source of information or actual co-ordination of relief activities during and after the floods. The identification of the interests and experience of such active individuals could lead to liaisons between the users and ER organisations to provide mutual benefit.

In terms of the temporal information, 787 references were extracted from 679 posts. 273 of these references were to specific times (e.g. “around 2.30pm”, “from 11-12 this morning”), 338 were more general relative references (e.g. “this morning”, “yesterday”) and the remaining 176 referenced dates (e.g. “Monday”, “on the 21<sup>st</sup>”). The most useful of these are the more specific times, as there is more potential to use this information to determine the start/end times of the events described in the posts. The majority of the specific times were contained in posts published on the main day of the flooding, which further increases their potential usefulness.

For the extraction of geospatial information a number of gazetteers were employed, which provided a very detailed geo-coding resource. Where an unrecognised frequently used word or geo-spatial expression was found, these were compared against the gazetteer using a Levenshtein distance spell-check to correct simple errors (e.g. Netheredge = Nether Edge, Bladwin Street = Baldwin Street). This approach is not effective at resolving the more creative colloquialisms, such as “Eccey Road” for “Eccleshall Road”. Table 1 shows the breakdown of extracted locations according to granularity levels.

Table.1 Extracted locations and number of references in posts.

	City	District	Street	Other	Total
<b>Locations</b>	5	69	197	40	311
<b>References</b>	281	585	473	348	1687

“Other” relates to locations such as villages or map features such as rivers or reservoirs. Location ambiguity was only found at the street level, aside from a village called “Fence”, a term which occurred 3 times in the posts. However the term “Fence” did not occur in conjunction with

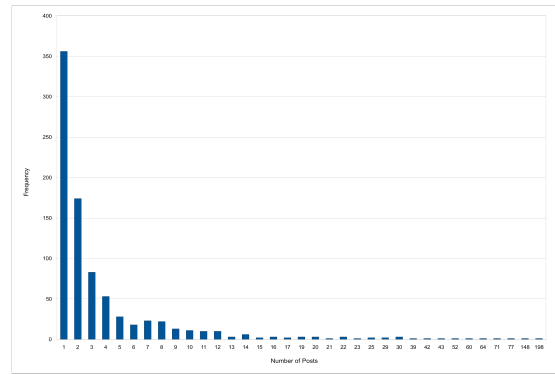


Fig.2 Frequency of user posting

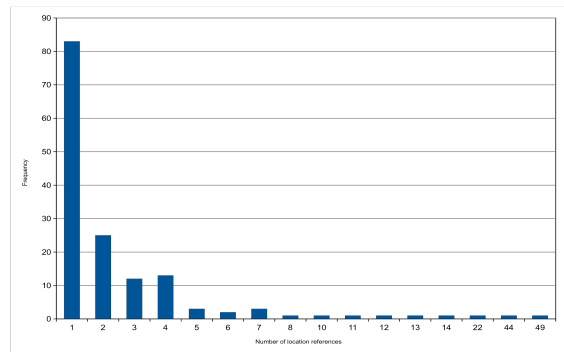


Fig.3 Frequency of location references in posts

other locations and upon examination of the text did not refer to the village.

Figure 3 shows the distribution of street name references. 83 names (42%) receive only a single reference, with only 10% of names referred to 5 times or more. Examining the names shows that there is some correlation between the number of references and importance of the street (as a thoroughfare) and the degree to which it was affected by flooding, with the two most frequent references being of an area which had to be evacuated and the main road into the north of Sheffield which collapse and took over a year to repair.

Of the 197 street names 97 were ambiguous and these received 224 references. Unfortunately it is not possible to do an exact evaluation of the ambiguity resolution as only the original poster could determine the true location. How-

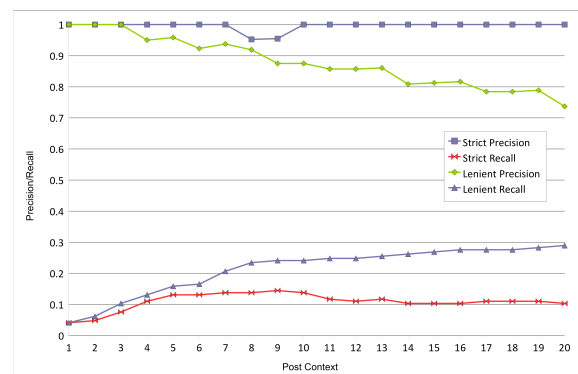


Fig.4 Precision/Recall of location resolution

Table.2 Topic cluster terms and interpretation

Cluster (stemmed) Terms	Interpretation
lightn thunder	General discussion of the intensity of the storm
pavilion towel bed	Concerning request for relief of citizens evacuated to The Pavilion
freezer power yedl postcod electr powercut tesco	Related to power-cuts, which occurred at different times in different postcodes. Concerns were expressed over food wastage. Note that YEDL is the electricity distributor for the area.
water manhol flow strand	Discussing CCTV footage of the floodwater pouring through a man-hole cover and stranding vehicles.
clousur flood close divert col-laps underwat tram gridlock	Concerning travel difficulties.
burst dam river trap evacu	Discussing the possibility of (Ulley) dam collapsing and its consequences.

ever, a Gold Standard was created, by determining the most likely location of the ambiguous cases. The location disambiguation process resolved 102 cases, producing 4 errors (i.e. Precision 96%, Recall 45.5%). Using other locations in the same post and directly quoted posts, produced 20 and 6 resolutions, respectively, with no errors. Similarly using locations from the same user and applying the strict approach to disambiguation, i.e. no contradictory information, also produced no errors, resolving 27 locations. However, applying a more lenient approach to the user locations doubled the recall (from 27 to 54) and produced only 3 errors (94.4% precision). Therefore, as would be intuitively expected, users and posts tend to refer to specific locations.

Figure 4 shows the precision/recall for resolving locations using proximate posts, applying a strict and lenient approach with a context of from 1 to 20 preceding posts. The maximum recall point (14.5%) of the lenient approach was deemed to be the best solution due to the high precision (95.5%).

The process of determining the most discriminatory terms in distinguishing flooding posts highlighted 33 (stemmed) terms. The hierarchical clustering technique was then applied to these, which formed the six clusters show in Table 2, with an accompanying interpretation of the cluster derived from examining the posts containing the cluster terms.

Four other terms (rain, scari, puddl and float) were not included in any of the clusters. The interpretation shows that meaningful clustering is possible. One caveat to express in relation to the clustering is that it was formed from all the flooding posts and therefore would only be available to interpret posts after the flood or could be used to build a topic model for future similar events.

The initial work of event identification used a strict but robust approach to defining an event, i.e. posts had to reference the same location (that is have a street location no

more than 100 metres apart) and contain co-occurring terms from the same topic-cluster. Only clusters that contained multiple users and five or more posts were considered, to ensure a degree of validation. The “lightning/thunder” topic was ignored. This resulted in the generation of 32 separate clusters, which contained 592 posts in total.

Examining the clusters shows that they do produce recognisable events, mainly related to power outages and street closures in given locations. The largest clusters, containing 146 posts, are related to the significant flooding and evacuation of a particular area (Winn Gardens) in Sheffield.

## V. RELATED WORK

Currently the most widely used systems employed to aid the ER Knowledge Management process deal, specifically, with interactions within and between ER organisations. However, there are also some initiatives exploring novel ways to aid citizen communication during an emergency<sup>6</sup>. Some of these systems, most notably Sahana<sup>7</sup>, provide a platform on which organisations can broadcast information to citizens and also provide a means for citizens to communicate with each other. The work described in this paper could be used to leverage the information collected on such platforms, potentially improving the ER process.

The general ideas behind extracting explicit useful information from individuals in public forums (or “the social web”) to form a collective intelligence are well discussed in [5]. In implementing these ideas a number of approaches have been applied to identify topics/themes/stories from public forum posts. The majority of approaches use the textual content ([4][9]) of the post, although post timings ([7]) and (user) social-networks ([3][10]) have also been exploited. Whilst the approaches use clustering techniques akin to the ones presented in this paper, the data sets are generally taken from well-established newsgroups and use a large collection of historical posts to construct the linguistic models.

The work on event tracking has generally been in the news domain (e.g. [1]), which tends to be well edited and have a relatively consistent format; as they are specifically designed to be easily read by the general public. However the general event tracking approach, linking of posts through named-entities and keywords, is the same as has been applied in this work.

One interesting area of work, which has not been explored, is the ranking of posts. This has been examined in [11] where posts are ranked in order to highlight the most relevant, original and informative posts. It would be interesting to extend this work to rank the relative importance of event clusters.

## VI. FUTURE WORK

The work presented in this paper is the initial step in trying to extract community information concerning an emergency situation from forum postings. The work will con-

<sup>6</sup> A comprehensive list of systems are described at <http://www.w3.org/2005/incubator/eiif/wiki/EMSystems>.

<sup>7</sup> <http://www.sahana.lk/>

tinue with a more comprehensive examination of the topic and event identification process. Examining how the temporal and user (social) network can be utilised to provide information context and determining how handle the process of continual event identification during the emergency when new posts (and the information they contain) may necessitate reformation of topics and events.

One issue which requires further work is to provide a more effective evaluation of the resultant events, i.e. how well does the extracted information reflect the actual events during the emergency. Current negotiations are underway to acquire the emergency services (Police, Fire, Ambulance) event logs, which will hopefully provide a reference upon which to evaluate the system. In addition it is hoped that they will be some gain from combining the organisational and community information, i.e. the community information system will add-value to the organisational information.

One key factor of the system not described in this paper is the ability to visualise the extracted information. Therefore a faceted browser is currently being implemented which allows the user to focus on specific values for the extracted information. In addition to selecting posts/events according to their time, location, user, content, etc., a timeline and map view is also available, facilitating the user in discovering patterns. As part of the visualisation system images and videos, which have been extracted from the posts, will be processed to determine the key images and then displayed. Within the WeKnowIt project, the User Requirements gathered from the ER personnel indicated that this visual information could be more informative and objective than a textual interpretation.

Another important consideration for the applicability of the work is to evaluate the impact of the quality of the geospatial resources. For the described case study high quality resources were employed, however these will not be available for all areas, especially if the emergency is in less well-developed countries. In the case study, the event identification used the most fine-grained locations; therefore the use of less precise or complete location information needs to be explored. However, with the development of mobile technologies and GPS and the increased availability of mobile broadband connection and mobile applications, more accurate and real-time information may also become available in the future.

Another area of work within the WeKnowIt project is examining ways to encourage the user to provide the information necessary to improve the collaborative knowledge. If the user is provided with an interface, which facilitates the provision of geospatial, temporal or topic information, or allows feedback on the automatic suggestions, then the quality of the event identification will be improved.

## VII. CONCLUSION

Despite the fact that the individuals who are posting to the forum have no over-arching coordination, and that they express information that relates solely their own experience and interest, it has been shown that the use of information

extraction, topic and event identification techniques can make explicit the collective information, which can provide a degree of global situational awareness.

Each post can provide some information about the situation, however due to the conversational nature of the posts these are likely to contain relatively low amounts of, potentially ambiguous, information. Therefore to utilise such social interaction data it is necessary to both extract relevant information from each post and to consider the context of that information. It is then the examination of the patterns produced by the combination of the related information that can provide a collective picture of the situation.

## VIII. ACKNOWLEDGEMENT

This work has been supported by the EU project WeKnowIt (ICT-215453).

## IX. REFERENCES

- [1] Best, C. and Pouliquen, B, Steinberger, R., van der Goot, E, Blackler, K, Fuat, F., Oellinger, T. and Ignat, C., "Towards automatic event tracking", Proceedings of IEEE International Conference on Intelligence and Security Informatics (ISI'2006), Lecture Notes in Computer Science, LNCS 3975, Springer-Verlag, Berlin Heidelberg, New York}, 26–34,
- [2] Comscore: "Leading User-Generated Content Sites See Exponential Growth in UK", 2006. <http://www.comscore.com/press/release.asp?press=993>.
- [3] Fortuna,, Blaz and Rodrigues,, Eduarda Mendes and Milic-Frayling,, Natasa. "Improving the classification of newsgroup messages through social network analysis", {CIKM '07: Proceedings of the sixteenth ACM conference on Conference on information and knowledge management}, 2007, pp 877–880
- [4] Glance, N., Hurst, M., Nigam, K., Siegler, M., Stockton, R., and Tomokiyo, T. "Deriving marketing intelligence from online discussion." In Proceedings of the Eleventh ACM SIGKDD international Conference on Knowledge Discovery in Data Mining (Chicago, Illinois, USA, August 21 - 24, 2005). KDD '05. ACM, New York, NY, 419-428. DOI= <http://doi.acm.org/10.1145/1081870.1081919>
- [5] Gruber, T. "Collective knowledge systems: Where the social web meets the semantic web." Journal of Web Semantics, 6(1):4–13, 2008
- [6] Nielsen: "User-Generated Content Drives Half of U.S. Top 10 Fastest Growing Web Brands", 2006. <http://www.nielsen-online.com/pr/PR060810.PDF>.
- [7] Qamra, A., Tseng, B., and Chang, E. Y. "Mining blog stories using community-based and temporal clustering". In Proceedings of the 15th ACM international Conference on information and Knowledge Management (Arlington, Virginia, USA, November 06 - 11, 2006). CIKM '06. ACM, New York, NY, 58-67.
- [8] Porter, M.F. "An algorithm for suffix stripping" Program, 14(3) 130–137, 1980
- [9] Spangler, WS, Kreulen, JT, Newswanger, JF "Machines in the conversation: Detecting themes and trends in informal communication streams." IBM Systems Journal, 2006 - [www-tr.watson.ibm.com](http://www-tr.watson.ibm.com)
- [10] Tuulos, V. H. and Tirri, H. "Combining Topic Models and Social Networks for Chat Data Mining". In Proceedings of the 2004 IEEE/WIC/ACM international Conference on Web intelligence (September 20 - 24, 2004). Web Intelligence. IEEE Computer Society, Washington, DC, 206-213.
- [11] Wanas, N., El-Saban, M., Ashour, H., and Ammar, W. "Automatic scoring of online discussion posts". In Proceeding of the 2nd ACM Workshop on information Credibility on the Web (Napa Valley, California, USA, October 30 - 30, 2008). WICOW '08. ACM, New York, NY, 19-26.
- [12] WeKnowIt. "D7.1 consumer and emergency response use case initial requirements", 2008. <http://www.weknowit.eu/system/files/D7.1.pdf>.
- [13] Winerman, L. "Crisis Communication" Nature 457(7228): 376-8.