

A Vision for Technology-Mediated Support for Public Participation & Assistance in Mass Emergencies & Disasters

Leysia Palen[†], Kenneth M. Anderson[†], Gloria Mark[‡], James Martin[†], Douglas Sicker[†], Martha Palmer[†], Dirk Grunwald[†]

[†]University of Colorado, Boulder
Department of Computer Science
430 UCB
Boulder CO 80309-0430

{palen,kena,martin,sicker,palmer,grunwald}@cs.colorado.edu

[‡]University of California, Irvine
Department of Informatics
Donald Bren Hall 5042
Irvine, CA 92697-3440
gmark@ics.uci.edu

We present a vision of the future of emergency management that better supports inclusion of activities and information from members of the public during disasters and mass emergency events. Such a vision relies on integration of multiple subfields of computer science, and a commitment to an understanding of the domain of application. It supports the hopes of a grid/cyberinfrastructure-enabled future that makes use of social software. However, in contrast to how emergency management is often understood, it aims to push beyond the idea of *monitoring* on-line activity, and instead focuses on an understudied but critical aspect of mass emergency response—the needs and roles of members of the public. By viewing the citizenry as a powerful, self-organizing, and collectively intelligent force, information and communication technology can play a transformational role in crisis. Critical topics for research and development include an understanding of the quantity and quality of information (and its continuous change) produced through computer-mediated communication during emergencies; mechanisms for ensuring trustworthiness and security of information; mechanisms for aligning informal and formal sources of information; and new applications of information extraction techniques.

Crisis Informatics, Disasters, Emergency Management, Natural Hazards, Social-Technical Systems, Social Computing, Technology-Mediated Social Participation, Visions of Computer Science

1. INTRODUCTION: A NEW VIEW OF DISASTER TECHNOLOGY

In an increasingly global society and on a planet experiencing effects of climatic change, large-scale emergencies both instigated by humans and arising from nature can devastate human life and our tightly-woven social fabric. With a promise of improved warning and coordination, a prevailing hope is that information and communication technology (ICT) can help reduce the impacts of large-scale disruptions, including political crises, natural disasters, pandemics, and terrorist threats. It is reassuring to know that interest in disaster technology development is growing and that the needs of first responders and emergency management agencies are receiving this attention (see for example, Landgren, 2006).

However, we argue that an opportunity to push computer science research in new directions is before us. We hope to grow a program and vision of research that focuses instead on an understudied but critical aspect of large-scale emergency response—the needs and roles of members of the public.¹ By viewing the citizenry as a *powerful, self-organizing, and collectively intelligent force*, ICT has the potential to play a remarkable and transformational role in the way society responds to mass emergencies and disaster. Furthermore, this view of a civil society that can be augmented by ICT is based on social and behavioral knowledge about how people truly respond in disaster, rather than on simplified and mythical portrayals of

¹ We use “citizens” and “members of the public” interchangeably to support readability, though the latter is the most accurate term. When readability allows, we privilege “members of the public” in acknowledgement that not all people who are affected by emergencies and other disruptions are accorded national citizenship. In fact, vulnerable populations, which include non-citizens, are disproportionately affected by such situations. Tourists and visitors to foreign countries are also affected by disruptions, and usually have fewer resources for reacting to events than residents.

people unable to help themselves (Fischer, 1998). Research has shown that disaster victims themselves are the true “first responders,” frequently acting on the basis of knowledge not available to officials (Stallings & Quarantelli, 1985; Aguirre et al., 1995; Kendra & Wachtendorf, 2006).

We argue that this transformative view is critical to our global future: When large-scale emergencies happen, there is often no way to survive it in practical terms unless we rely on each other for help. The urgency and scale of many disaster events are such that no one—not even the most experienced and best technology-equipped responders—can rescue all victims or direct all people over the span of the event as to what the best course of action might be. Climatic change and population migration to geographically vulnerable areas mean that naturally occurring hazards will exert increasingly extensive damage. Man-made and terrorist threats can also have greater potential to cause lasting damage to the social and built environment. The practical answers to responding to these scaling needs cannot only be to hire increasingly more first responders. It is instead necessary, through innovative ICT, to leverage the power of the collective intelligence of the citizenry to support natural instincts, which are to *search for reliable information using any means possible to optimize for local conditions* (Sorenson & Sorenson, 2006). To this end, we propose a vision of research that makes use of and enhances information supplied by members of the public and official sources, and supports members of the public to become better “everyday analysts” with application services that support their immediate needs in times of disaster and mass emergency.

This view challenges *fundamental assumptions about how information should be controlled and disseminated to ensure public safety*. That officials can and should provide the best information during emergencies, and that the public can and should primarily rely on official information appears to be a difficult-to-resist, near-universal human hope. However, it is critical that we adopt a view of broad public participation in emergency response as soon as possible.

Technology is already expanding peer-to-peer information-seeking behavior through *computer-mediated communication* (CMC) in remarkable ways through community websites, email, broadcast texting, and other media. Through ICT, people are “going local” to tap the knowledge of their fellow residents to make critical decisions. In research during the 2007 Southern California fires (Sutton, Palen, Shklovski, 2008; Shklovski, Palen, Sutton, 2008), unofficial, ICT-enabled “backchannel” communications activity supported self-help as described by research participants:

“The only way we all have to get good information here is for those who have it to share it. We relied on

others to give us updates when they had info and we do the same for others.”

“What we learned...is that there is no ‘they.’ ‘They’ won’t tell us if there is danger, ‘they’ aren’t coming to help, and ‘they’ won’t correct bad information. We regular folks have to do that amongst ourselves.”

In this event, there was simply too much rapidly changing activity over an expansive area for officials and the news media to capably support the vast and highly localized information needs of the public.

New Vision, New Objective

This vision for a new future in emergency response is being realized in part by research our team is conducting in an effort we call *Project EPIC: Empowering the Public with Information in Crisis* (Project EPIC, 2009), funded at the level of US\$2.8M by the US National Science Foundation. Our hope is that this research project contributes solutions for short-term needs as well as a long-term vision in the emergency space through applied and basic research.

The objective for such a vision is that people find and share the information they need to make the best, most informed decisions possible during situations of emergency and high disruption—in other words, to *support and enhance the analytical skills of members of the public*. In our own work, we aim to do this through leveraging the knowledge of members of the public, and through reframing *emergency response as a socially-distributed information system*. This work will support people across a range of geographies (Powell, 1954), by which we broadly mean those who are inside the crisis zone as well as those on the outside who are compelled to help not only through supplying physical labor, goods and psychological support, but also by playing critical roles in information dissemination. Through reuse of computer-mediated communications that are publicly available (such as those on community websites, personal blogs, public texting systems like Twitter, social networking sites, mapping sites, etc.), we aim to derive applications and services for use by the public that integrate those original communications with information that helps citizens assess context, validity, source, credibility, and timeliness to make the best decisions for their highly localized, changing conditions.

Such a vision of research, then, is guided by this question: *How can publicly-available, grassroots, peer-generated information be deemed to be trustworthy, secure and accurate, so that it can be leveraged and aligned with official information sources for optimal, local decision-making by members of the public?*

The answer to this question requires programmatic vision to gain significant ground as soon as possible, but it will nevertheless remain a core and even guiding challenge for our future’s big computing challenges that

we must face today. This research-based, empirically-supported integration, extraction and presentation of information to directly support members of the public during crisis is based on a significant shift in perspective of emergency management as a socially-distributed information system. Imagine this scenario of public activity and local decision-making in the not-too-distant future:

It is Fall 2014 in Southern California. Temperatures are high and small fires have ignited in dry areas. With Santa Ana winds picking up speed and fuel for the fire—dried vegetation—at its densest in years, emergency officials are worried that the fires will expand and affect communities that have encroached further and further into the “wildland-urban interface.” Residents have been notified of the fires through weather and environmental alerting directly to their smart phones. Some evacuations have begun, and residents in other more distant areas remain on alert.

Upon receiving the evacuation and alert notifications, people start taking action. Most residents under evacuation prepare to leave quickly, and use smart phone applications to get information about evacuation routes and traffic conditions. In addition, they post about their plans to leave; where they plan to go; as well as requests for help that they might need once they evacuate, including places to stay, medical assistance for pre-existing health concerns, and help with children and pets. This information gets propagated to friends, family and neighbors, as well as officials who can then make better plans about who might still need evacuation assistance. Neighbors, all concerned about their communities, make plans to stay in close digital coordination to pool information they are able to gather about what is happening in their area.

Others in more distant areas are also engaged in intensified information search in case they need to evacuate quickly. They go online to get specific information about the fires from a wide range of peer sources who know well the particular areas under threat. After all, some areas are still rural, and require local knowledge to convey precise information about what is happening and what neighbors down the road are doing. Rather than attempt to assemble this online information by hand and check it against heavily taxed official websites with possibly aged information as in the past, they now have access to a suite of synthesized, easy-to-use web applications tailored to crisis situations that integrate up-to-the minute information from multiple sources. No longer must they worry about not knowing what new technology is available and missing critical information that resides in those new forums. Instead, they can readily find the most relevant online destinations and information streams with supplemental meta-information about how trustworthy, timely and relevant the information is to their particular circumstances.

Multidisciplinary Orientation

Realizing this vision of transparent, trustworthy information access, integration and delivery requires multidisciplinary research specialized in the area of emergency and crisis conditions. Empirical research with analysis that directly inputs into information extraction and data mining techniques will make the solutions “ecologically valid” and practical. Solutions in this social space are different than they are for other domains running under so-called “normal” conditions. A program of research in this area poses a set of research questions that integrates empirical research and innovation, and that (at least) includes expertise in software engineering, information science, human-computer interaction, information extraction, natural language processing, network security, telecom policy, and emergency management.

2. CRISIS INFORMATICS: PERSPECTIVE, RATIONALE & SUPPORTING RESEARCH

We place this research vision into a broader frame of investigative engagement known as *crisis informatics* (Hagar & Haythornthwaite, 2005; Palen, et al. 2009). Our view is that *crisis informatics* addresses socio-technical concerns in large-scale emergency response. Additionally it expands consideration to include not only official responders (who tend to be the focus in policy and technology-focused matters), but *also* members of the public. It therefore views emergency response as a social system where information is disseminated within and between official and public channels and entities. Crisis informatics wrestles with methodological concerns as it strives to develop new theory and support informed development of ICT and policy (Palen, 2009).

The simplified schematic in Figure 1 illustrates our conceptualization of changes occurring within the information arena of disaster, and the pressures these changes place upon the “interface” between the formal and informal response. The white arrow (A) denotes conventional management of public information as unidirectional from the formal to the informal spheres. This assumption of unilateral communication—though never accurate in practice—is the basis for US local and federal policy decisions, and is similar to other nations’ emergency management arrangements. The dark gray arrow (B) illustrates that existing peer processes of information creation and dispersion is now expedited and proliferating because of pervasive ICT. This, in turn, has two consequences of concern denoted by the black arrows: these changing peer communications now spawn tactical data and human or crowd convergence information that is often of interest to the official response (C) both to leverage and control.

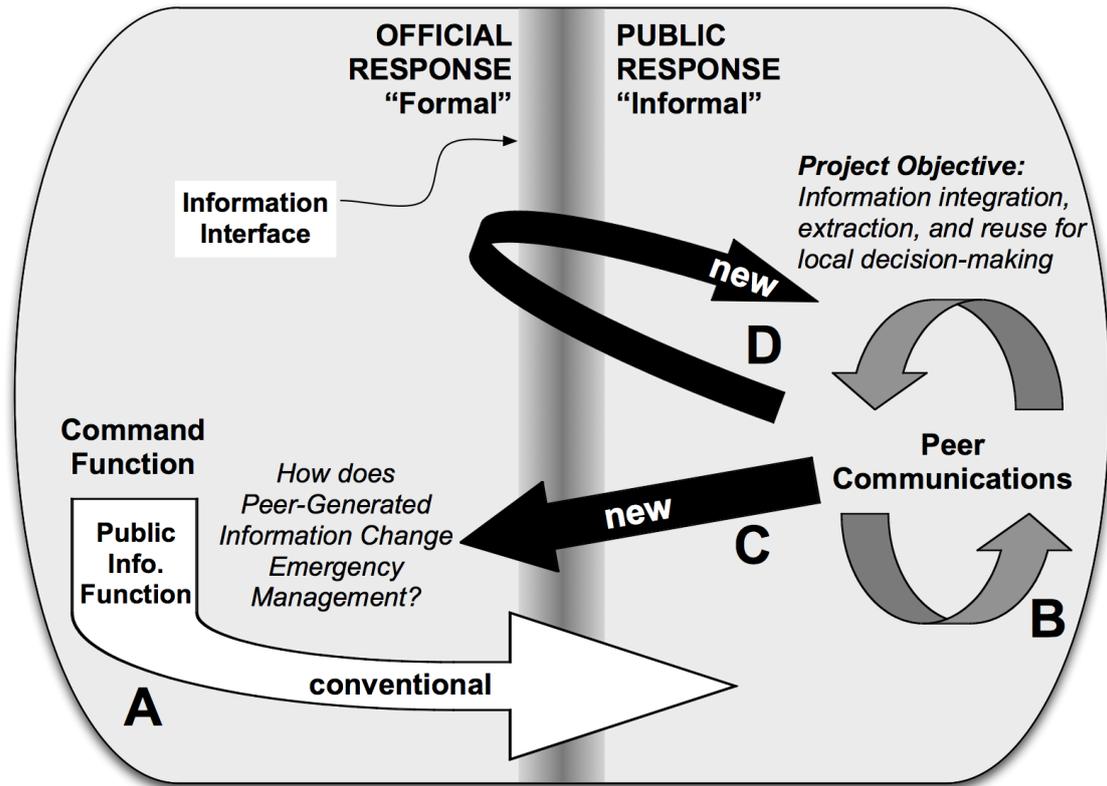


Figure 1: Crisis Response as a Social System: High-Level Schematic of Conventional and New, ICT-Enabled Information Pathways

The new issue depicted by arrow C deals with the matter of where this new information is received in the organization of emergency response: Are these data that result from communications something that the “public information” function attends to, or does the “intelligence” arm of command field them? That we can ask this question speaks to the new tensions that are arising in the institution of emergency management, and is requiring it to reconsider how information is shared and received. This argument of “changing information pathways” is one we have made in previous papers (e.g., Palen & Liu, 2007) and have presented to practitioner audiences. This issue is actively being reviewed by government agencies in the US and groups, including the US Federal Emergency Management Agency, where the new presidential administrative appointee, Craig Fugate, is now calling for increased support of such change (O’Keefe, 2009).

Research on Citizen Activity

The rationale for such a vision of computing research is based on knowledge of existing social behavior in crisis events (Dynes, 1970; Mileti, Drabek & Haas, 1975; Drabek, 1986; Tierney, Lindell & Perry, 2001). High involvement by members of the public in disaster is not new, though ICT makes their role more visible and broadens the scope of their participation. “First responders” are not, in practice, the trained professionals who are deployed to a scene in spite of

the common use of that term for them; they are instead people from the local and surrounding communities (Dynes, 1970) who provide first aid, transport victims to hospitals in their own cars, and begin search and rescue (Fischer, 1998; Tierney, Lindell & Perry, 2001). Indeed, in cases of structural collapse, the majority of those saved are by local emergent volunteer groups (Aguirre, 1993). Individuals and local groups come together as emergent temporary organizations that *improvise* rescue and relief efforts (Mendonça & Wallace, 2004; Wachtendorf & Kendra, 2005). For example, following 9/11, ferry captains and others self-organized to systematically evacuate people from Manhattan Island (Kendra & Wachtendorf, 2003). Members of the public in the immediate and surrounding area of a disaster provide relief in the form of shelter, clothing, blankets and food (Tierney, Lindell & Perry, 2001). Victims of disasters—to the extent they are able—are themselves often among the volunteers. It is also the local community that manages and shapes long-term recovery once the immediate response effort is complete. The public is central to disaster response.

In addition, people are natural information seekers, and will seek information from multiple sources, relying primarily on their own social networks—friends and family—to validate and interpret information coming from formal sources, and then to calculate their own response measures, such as when they will leave and

under what conditions. Several studies (e.g., Drabek et al., 1975) found that people will collectively decide with their extended family, friends and neighbors their evacuation behavior based on their interpretation of warning content, its mode of communication, and its source.

ICT & Public Participation in Disaster Response

ICT is rapidly expanding this existing self-organizing, improvisational, information-seeking, and innovation behavior by members of the public. Importantly, ICT-enabled public self-organization around the *decomposition of problems* that present themselves in disaster response provides the means for *highly parallel and distributed problem solving*.

The ramp up to this point has been steep. Following the September 11, 2001 World Trade Center and Pentagon attacks during the disaster phase period that is characterized by intensive information search, people turned to the Web (Schneider & Foot, 2002), a time when blogs were not yet popularized. During this same time period, Hagar and Haythornwaite examined how farmers used computational media to find information and support one another during the lengthy 2001 UK farming crisis (2005; 2009). Indications of significant socio-technical change in public participation behaviors as mobile phones, personal blogging, and on-line photo repositories like Flickr became available (Liu et al., 2008) were notably apparent in the December 2004 Indian Ocean tsunami. In the 2005 London terrorist attack, camera phone users on subway trains offered officials, the media, and the public, first-hand assessments of the impact of the bombings (though this transmission of information was not necessarily real-time). Numerous public-initiated information sources and new forms of personal ICT use sprang up in the aftermath of Hurricane Katrina, including some of the earliest housing aggregator and giving sites, people-finding activity, and map-based mashups (Palen & Liu, 2007; Torrey et al., 2007; Shklovski, Palen, Sutton, 2008).

The tragic Virginia Tech shootings in April 2007 saw the emergence of social networking sites as destinations for collective disaster-related sensemaking, as students and others collected information on details of the shootings and reported on their own safety (Vieweg et al. 2008). The resulting interactions became the basis for the emergence of *collective intelligence* (Hiltz & Turoff, 1993; Palen, Hiltz, Liu, 2007) or socially distributed cognition activity (Hutchins, 1995) that “discovered,” in parallel and with redundancy and accuracy, the names of the 32 fatalities in advance of official releases of that information (Palen, et al., 2009). With the August 2007 Minneapolis Bridge collapse came new public attention placed on the possibilities of micro-blog messaging activity (Twitter) among members of the public, and increasingly, among traditional news media.

Similar activities over a more protracted time period during the 2007 Southern California wildfires showed the centrality of ICT-enabled community information resources and other “backchannel” communications (Sutton, Palen, Shklovski, 2008; Shklovski, Palen, Sutton, 2008). The use of social networking sites in the aftermath of the shootings at Northern Illinois University in February 2008 once again demonstrated how people engage in *widescale interaction* as they did during the Virginia Tech event, though with more apparent caution because of the awareness students had about being on the public, digital stage (Palen & Vieweg, 2008).

After the May 12, 2008 Sichuan earthquake in China, a popular internet forum became a location for integrating information from multiple sources, organizing public action and expressing grief and anger (Qu et al., 2009). During the Red River Floods in the US and Canada, and the Oklahoma grassfires (US), extensive research of twitter records examined the relationship between geographical location and the generation of new information to the discussion space (Starbird et al., 2010; Vieweg et al., 2010; Starbird & Palen, 2010). Analysis of the “neogeographers” who emerge during crisis events to create crisis map mashups reveals the relationship between design decisions and mashup evolution (Liu & Palen, 2010).

Development work to build upon ICT-supported public activity is gaining ground. Work in the area of humanitarian crisis, specifically the Kenyan post-election violence in January 2008, was the basis for the creation of a “crowd sourcing” environment, “Ushahidi,” where people could warn others and report violence. Initial findings on an analysis of the use of that system show that citizen journalism had a greater geographical reach than traditional sources but that use fell off quickly in this early instantiation of the kinds of behaviors we might envision as more common for the future (Meier & Brodock, 2008). Work at the University of Maryland on “community response grids” seeks to similarly leverage public participation in emergency response (Wu et al, 2008b; Shneiderman & Preece, 2007), and also included empirical examination of the adoption of college campus SMS early warning systems by students, which still remains lower than hoped (Wu et al., 2008a).

In the area of public health, disease trajectories are being tracked and anticipated through Internet search. Multiple efforts, including Google Flutrends, track flu outbreaks by analyzing and geolocating relevant query terms (e.g., “aches”). The US Geological Association (USGS) uses citizen geo-reports to measure the intensity of earthquakes on “Did You Feel It” (<http://earthquake.usgs.gov/eqcenter/dyfi/>). The USGS has also deployed software using common accelerometer-based laptops to detect earthquakes, and is now mining twitter for earthquake detection. This again suggests a future vision of computer-augmented

users as participatory data collectors. Finally, examination of personal ICT use by Israeli and Iraqi citizens suggests that it can help people repair broken daily routines and substitute for face-to-face social interaction that is severed during ongoing, indefinite wartime disruption (Mark & Semaan, 2008). Blogs, Internet forums, and email distribution lists enabled them to communicate with people across the globe residing in undisrupted environments, which helped restore a sense of normalcy to their lives (Mark et al., 2009; Al-Ani, Mark, Semaan, 2010). Finally, the Information Systems for Crisis Management and Response organization and conference is a forum where new issues in ICT and crisis are being explored (ISCRAM, 2009).

3. A RESEARCH PROGRAM

We envision a future where members of the public can get information about an emergency that is specialized to their needs—as well as meta-information specialized for crisis situations that helps them make judgments about the ever-growing amount of information. Such meta-information includes features of its source; judgments about the authoritativeness of the source; its concurrence (or not) with official sources; its timeliness (as the information may be *better* than official sources); other spatio-temporal features of the information's life; and consideration of who would be looking for this information and why. The discovery of what kind of meta-information is necessary will need to be derived from ongoing empirical research, which our group is conducting, but for which there remains additional need by complementary efforts. Data mining and information extraction techniques have a critical place here, as does the creation of trust models and security techniques. The combination of empirical grounding and computing innovation is required to achieve our goal of *scaffolding the veridicality of peer-generated information with supplemental, computationally-derived meta-information that lends additional context for better, local decision-making*.

Setting Initial Directions: Research Questions

This vision enumerates a number of questions. We outline five primary questions that we are addressing.

Empirical Research on Technology-Mediated Behavior

RQ1: *What kind of information is being generated and distributed through computational media by members of the public in times of crisis? How do people self-organize around this information? How do they assess accuracy and trustworthiness?*

Extending traditional “quick response” disaster research strategies (Stallings, 2002), research on the study of CMC before, during and after events is a new terrain that requires adaptation of current methods to tackle such a massive number of data points. The

problem here is never quite knowing what the universe of interaction is, and therefore how to properly sample it, and selectively choose from it. It is here where we first see that to tackle the challenge of technology-mediated social participation is to tackle what many are doing in the new E-Social Science field (see Halfpenny & Procter, 2009) as a primary example among others. Early work in this area was known as digital or virtual ethnography (Jones, 1999; Miller & Slater, 2000; Hine, 2000, 2005), though we must already modernize this methodological orientation to depict real-time search and visualizations of massive data sets that support quantitative and qualitative analyses.

In tackling such empirical goals for disaster research, we must understand features of activity at different scales of interaction. This includes (but is not limited to) analysis of:

- (i) where and when online crisis-based information arises;
- (ii) how people organize around online information exchange including the degrees and types of participation;
- (iii) who participates and why;
- (iv) the presence or absence of self-policing for accuracy and credibility;
- (v) the management of timeliness of information;
- (vi) the kind of information needs that arise for different kinds of large-scale emergency events, and who needs what kinds;
- (vii) how people make judgments about the accuracy/validity of information from peers.

From this, we need to engage in analyses to characterize the kinds of information and social roles that people take to generate requirements for innovation. This research is producing analyses that describe such behavior, particularly with respect to microblogging activity, (Starbird et al, 2010; Vieweg et al, 2010; Starbird & Palen, 2010), and is supporting the development of information integration and data mining techniques for accurate, informative characterization of computer-mediated, self-organizing behavior.

Finally, engagement with emergency response practitioners, including the use of participatory design techniques (Greenbaum & Kyng, 1992), is critical here. Though the vision of the work is driven by an overarching goal of supporting the needs of the public rather than monitoring them, these same goals of communication of good information are espoused in the work of emergency personnel, particularly the “public information officers” who liaison with members of the public to support accurate dispersion of information. It is critical that our modelling of behavior that we see on-line both informs and is informed by established and perhaps changing practice of practitioners.

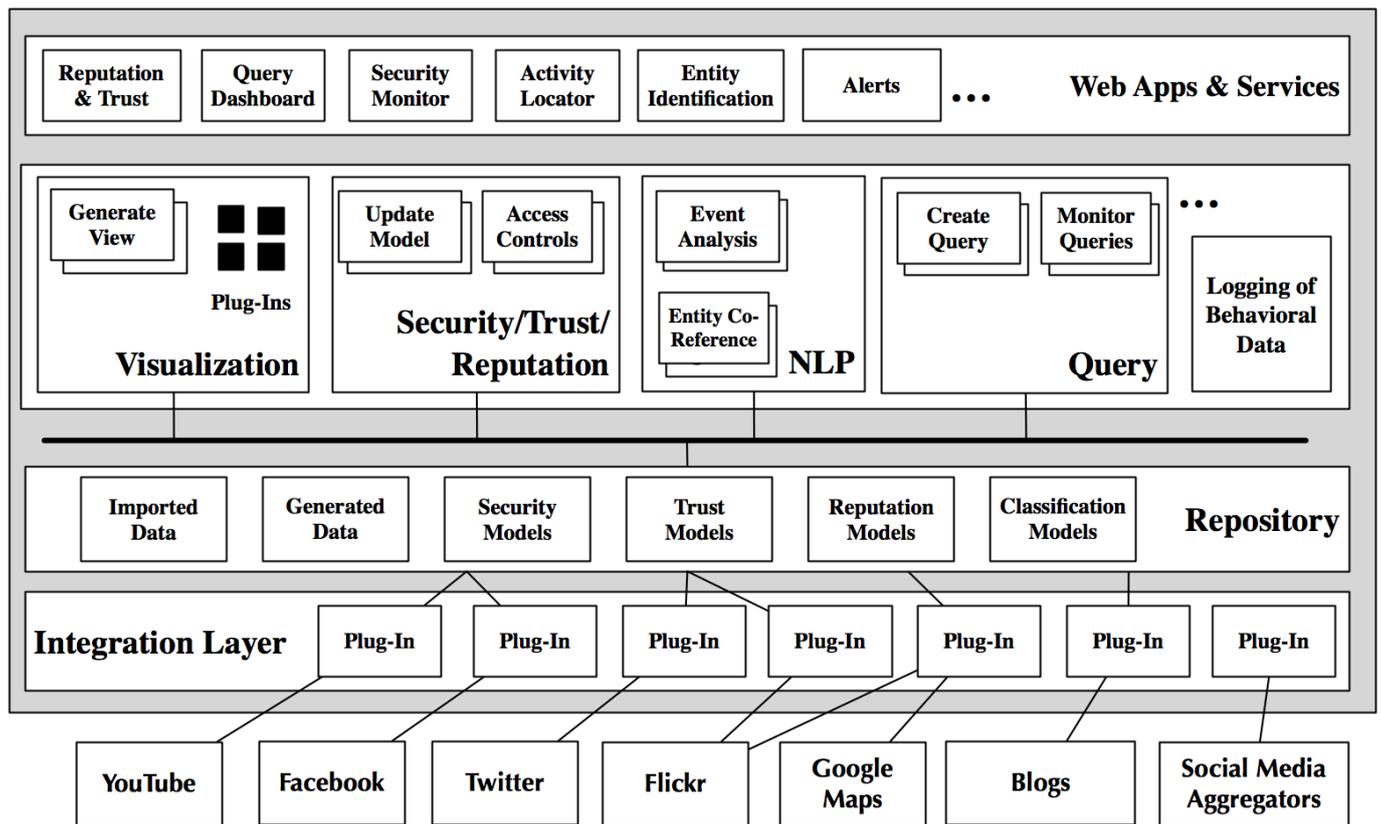


Figure 2: The Information Integration Landscape. Members of the public provide information about crisis events via social media. This information is analyzed/vetted and then delivered back to the public in annotated form via Web-based software and services.

Information Integration

RQ2: How can the extreme heterogeneity of information generated during crisis events be tamed such that it can be transformed into useful knowledge for citizens?

We see *information integration* (Anderson et al., 2003) as a core computational research concern of this crisis informatics-informed technology space. During the emergency period of disasters, vast heterogeneous information is produced from multiple sources. New techniques and tools are needed to usefully process this information in an automated fashion. The basic software architecture (Figure 2) therefore sits on top of the rich information landscape of Web 2.0 sites and services and feeds key aspects of that information through the analytical tools that need to be developed in the areas of NLP, security, privacy and data validation. It then generates new information and services that the public can use to make decisions in a crisis event while also saving that information for later empirical analysis. This architecture assumes the use of scalable data integration services that already exist for popular social media/networking sites as well as Web 2.0 sites.

The integration layer of the software architecture populates a shared repository for other tools and services that produce meta-information associated with the original data that identify important entities, classify contributors according to the roles they enact and assume, maintain privacy for members of the public as their information is aggregated into summaries of an event's status, and annotate information with indicators as to its trustworthiness, accuracy and reliability (within appropriate timescales). In addition, visualization and mash-up services are needed that allow data to be merged in multiple ways for delivery to members of the public through a suite of Web-based applications and services optimized for mobile and desktop devices. These services will include applications that point users to active on-line forums, help them make sense of relative terms in posts via association with known entities surrounding an event, and visualize results of security and trust models.

Information Extraction & Natural Language Processing

RQ3: How can information extraction and in particular natural language processing techniques be extended to accommodate the emergence of different forms of crisis-related communication in real time, and what features of those communications will be most relevant for characterizing important elements in different kinds of events?

Semantic Mark-Up:

[Thousand Trails _{LOC}] will be open for [re-entry _{EVENT}] at [8am today _{TMP}]. Re-entry point will be [south _{GEO}] from [Chula Vista _{LOC}] at [Otay Road _{LOC}].

ReOpening event:

Locale: *Thousand Trails*

Time: *8am*

ReEntry Point: *Otay Rd*

Figure 3: Information Extraction on Microblog Data

The diversity and sheer volume of information sources generated during crises events pose major challenges for current NLP-based information extraction (IE) techniques. In our work, we are attempting to use NLP-based techniques to critically inform the development of classification services by use of robust and efficient IE. These services annotate language-based information streams from publicly-generated CMC sources with data that reveal the types of events and entities represented in those streams, the roles of those entities in the events in question, and any temporal and geospatial information that can be gleaned. These annotations will serve as critical inputs for determining relevance of the information to different members of the public. They aim to facilitate the detection of new and anomalous events, enable comparison of information both across sources and crisis, and facilitate data mining of CMC information streams.

Consider the excerpt in Figure 3 from the text-based Twitter stream generated by a local resident turned “information broker” during the 2007 Southern California fires: “*Thousand Trails will be open for re-entry at 8am today. Re-entry point will be south from Chula Vista at Otay Road.*” As the following semantic markup in Figure 3 illustrates, this example contains three named locations (two communities and 1 road), one temporal relation and one geospatial relation, as well as a future event.

Previous research in IE has shown that identification and indexing of basic elements such as these can facilitate the creation of systems that can reliably detect and classify more complex recurring events of interest to the general public. For example, the annotations shown in Figure 3 might be used to instantiate and populate a “reopening” event frame as shown. Further temporal and geospatial processing can normalize the mentioned times and location to specific times and locales. The identification and classification of recurring events such as the opening of shelters or closure of roads will facilitate the creation of services that allow users to create ad hoc or standing requests for alerts about types of events, or events occurring at or near particular locations.

Current NLP-based approaches to information extraction are trained using supervised machine learning techniques to reproduce the sets of annotations produced by human analysts (Jurafsky & Martin, 2008). While the bulk of this work has been applied to edited newswire-style texts, some of the most exciting and challenging areas of research involve the application of these techniques to new forms of social media that serve as the channels for the kind of user-generated content we see in crisis warning and response. The primary research challenge facing the application of IE techniques here lies in translating the use of machine learning techniques that are trained on materials from previous events to the real-time emergent conditions and interactions during crisis situations. Our approach to this includes applying the annotation described above to our social media data collections, as well as using bootstrapping techniques to quickly detect key novel elements in new situations.

Information Security and Reputation Systems

RQ4: *How can diverse information security techniques be used to address such issues as privacy of participants, validity of data and appropriateness of information aggregation methods? How might different security methods conflict and potentially compromise one another?*

We are investigating the use of established and potentially new techniques, including information aggregation, privacy protection (Sicker, 2003), location privacy (Gruteser & Grunwald, 2003), confidentiality, reputation validation (Sicker & McCoy, 2007) and anonymization (Bauer et al., 2007), as they apply to the matters of security, privacy and accuracy in self-organization and peer communication in crisis events. The cause and effects of these security and privacy techniques are inter-related and interdependent: As some techniques strengthen some data qualities, they will weaken others. Research in this area therefore needs to identify and develop the need for additional methods that can balance issues such as privacy and reputation.

There are several security and privacy concerns within the proposed architecture. In particular, since users are sharing their perspectives and contributing information about a dynamic situation, it is necessary to consider user privacy, because such citizen reporting may put people at risk, particularly in places of ongoing disruption from political crises. An even more fundamental concern stems directly from the collaborative and peer nature of the information sharing: Malicious users may have an incentive to contribute false or misleading information, particularly in a situation of terrorism. If enough malicious users exist under terrorist or cyber-terrorist conditions, it may be possible to disseminate false information to others and disrupt natural social processes of self-policing

and accuracy ascertainment. This is a challenging but critical situation to address.

Due to the peer and collaborative nature of this system, it is possible to pool perspectives of several users and extract relevant information about the disaster in part by using ratings to rank information accuracy with Bayesian filtering to determine reputation from those ratings (Mui & Mohtashemi, 2002; Whitby et al., 2004). Though the user-provided information and rankings approach offers tremendous advantages such as fast information collection and aggregation, it is also necessary to provide certain privacy protections for contributors. For instance, users may not wish to share their physical location with others, in part because they may be violating official laws to leave certain areas, so mechanisms must be in place to offer privacy protection. Physical location is just one example of a weak identifier (i.e., pieces of information that can be used to identify individual users). Other identifiers include network addresses. A wealth of literature describes many techniques to effectively remove these network identifiers and achieve a degree of anonymity (Chaum, 1988; Reiter & Rubin, 1998; Dingedine et al., 2004). However, in our context, it may be more appropriate to allow users to have a high degree of control and flexibility with respect to the types of personally identifying information revealed.

However, if mechanisms are provided to enhance the users' privacy and even offer the capability for anonymous information contributions, it is possible that malicious entities may exploit this anonymity and inject false and misleading information into information streams. Fundamentally, establishing identity within a networked environment is a challenging problem. For instance, a classic problem within the field of reputation systems is the Sybil attack, which occurs when a single entity establishes a large number of identities, for instance, by adding a large number of nodes to a peer-to-peer network (Douceur, 2002). This has implications for designing reputation systems to detect the injection of bad information by malicious parties, since it is always possible for them to simply add a new identity to the system to counteract any previous misbehaviour.

A number of reputation systems have been proposed to combat against attacks such as the Sybil attack. However, these systems, considered "explicit" reputation systems, rely on quantitatively measurable activities or transactions (largely suited to e-commerce) and maintain databases of explicit feedback given by users—features that do not exist in typical social networks.

Given this fact, we take an alternate approach to determining reputable users during an event. We will look at two possible implicit reputation indicators, one using client location services and another using analysis methods from varying disciplines such as computer science, physics, biology and others that are

based on network graphs and activity within networks. Using client location services, it may be possible to determine location based on usage of the cellular or wireless 802.11 wireless infrastructure. However, access to this client information is dependent on obtaining information from service providers or users making location information available from their client programs. Using network analysis methods, it may be possible to infer location, influence, and other attributes based on communication, context, and association with others in the network. Data for these methods are available via public application programming interfaces (APIs) from some online social networks. By monitoring the APIs during crisis situations, we are able to use these analysis methods to infer information about users as the crisis is evolving.

Public Policy

The synthesis of these activities within our own group—but that we hope will be complemented by efforts elsewhere in this important area—aims to generate the basis for supporting a variety of tools and services aimed at supporting the citizenry during times of disasters and mass emergencies. Consequently, we must consider public policy issues that arise from this research and innovation:

RQ5: What public policy issues emerge as a result of this vision of empirical research and development activities, and how can they be framed using a crisis informatics lens?

Investigations of this kind require consideration about how crisis-related data collection, analysis, and dissemination within the context of current federal and state laws, policies, and regulations (Ohm, Sicker & Grunwald, 2007). Because information aggregation efforts may cross national borders, we will also take into account similar (and often more stringent) laws and policies in other countries. Important questions in this area center on issues of legal precedent, communications-related laws and policies, regulations and authorities regarding information dissemination during disasters, along with privacy and security issues. Other questions include how the policy context will affect our own information capture and aggregation methods. As an over-arching concern, we question how it will be possible to design methods for capturing and processing diverse content ethically and in adherence of law, while also not creating new, insurmountable policy issues around information aggregation (Ohm, Sicker & Grunwald, 2007).

CONCLUSION

This initial programmatic research vision for computer science focuses on an understudied aspect of large-scale emergency response—the information-producing and sharing activities by members of the public as supported by ICT. Intellectual contributions include the

expansion of broad thinking about societal-scale interaction in the area of crisis informatics, and bridging information science, cognitive science and computer science. They include the application of these empirical and policy implications to challenging matters of information integration and extraction, as well as trust modeling of highly heterogeneous data for transformation to useful knowledge. Innovation will be based on empirical study of computer mediated communication and citizen information needs in a range of crisis events, with the development of methods to analyze such behavior, and in light of privacy, security, ethical and public policy issues.

The initial directions here are based on a foundation of supporting work that has led to the very paradigm-shifting perspective we propose: that innovation for emergencies could greatly benefit by reframing disaster response as a set of socially-distributed information activities that support powerful, parallel, socio-technical processing of problems in times of change and disruption. Good quality information and meta-information that indicates accuracy and trustworthiness is what people need to make local decisions, to gain situational awareness and build resilience in the face of threat.

At its most ambitious, such a programmatic vision speaks to the difference between a society that painfully comes to this understanding after enduring repeated disasters, and one in which mechanisms for leveraging the collective intelligence of the public are accelerated in a systematic fashion, and with serious consideration to ethics, the practical aspects of emergency management, and human needs.

ACKNOWLEDGEMENTS

We thank the US National Science Foundation for their support of this research through grants IIS-0546315 and IIS-0910586. We thank the team associated with this research: Ban Al-Ani, Mossaab Bagdouri, William Corvey, Amanda Hughes, Sophia B. Liu, Matt Novinger, Travis Rood, Chris Schenk, Aaron Schram, Bryan Semaan, Catharine (Kate) Starbird, Sudha Verma and Sarah Vieweg; and advisors Ed Hutchins and Morten Kyng. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the US National Science Foundation.

REFERENCES

Aguirre, B. (1993). *Review of Social Science Research on Warning, Evacuation and Search and Rescue*. Hazard and Recovery Center, Texas A&M.

Aguirre, B., Wenger, D., Glass, T., Diaz-Murillo, M. & Vigo, G. (1995). The Social Organization of Search and Rescue: Evidence from the Guadalajara Gasoline

Explosion. *International Journal of Mass Emergencies and Disasters*, 13:93–106.

Anderson, K., Andersen, A., Wadhvani, N. & Bartolo, L. (2003). Metis: Lightweight, Flexible and Web-based Workflow Services for Digital Libraries. In *Proc. 2003 Joint Conf. Digital Libraries*, pp. 98–109.

Bauer, K., McCoy, D., Grunwald, D., Kohno, T. & Sicker, D. (2007). Low-Resource Routing Attacks Against Tor. In *Proc. 2007 ACM Workshop on Privacy in the Electronic Society*, pp. 11–20.

Chaum, D. (1988). The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability. *Journal of Cryptology*, pp. 65-75.

Dingledine, R., Mathewson, N. & Syverson, P. (2004). Tor: The Second-Generation Onion Router. In *Proceedings of the 13th USENIX Security Symposium*.

Douceur, J. R. (2002). The Sybil Attack. In *Revised Papers From the First Int'l. Workshop on Peer-To-Peer Systems*, Druschel, P. Kaashoek, M. & Rowstron, A. (Eds.) *Lecture Notes In Computer Science*, vol. 2429. Springer-Verlag, London, pp. 251-260.

Drabek, T. (1986). *Human System Responses to Disasters: An Inventory of Sociological Findings*. Springer-Verlag.

Drabek, T., Key, W., Erickson, P. & Crowe, J. (1975). The Impact of Disaster on Kin Relationships. *Journal of Marriage and the Family*, pp. 481–94.

Dynes, R. R. (1970). *Organized Behavior in Disaster*. Heath.

Fischer III, H. (1998). *Response to Disaster: Fact Versus Fiction & Perpetuation*, 2nd ed., University Press of America.

Greenbaum, J. & Kyng, M. (Eds.) (1992). *Design at Work: Cooperative Design of Computer Systems*. Lawrence Erlbaum Associates, Inc.

Gruteser, M. & Grunwald, D. (2003). Anonymous Usage of Location-based Services through Spatial and Temporal Cloaking In *Proceedings of the 2nd International Conference on Mobile Systems, Applications and Services*, pp. 25–32.

Hagar, C. (2009). The Information and Social Needs of Cumbrian Farmers during the UK 2001 Foot and Mouth Disease Outbreak and the Role of Information and Communication Technologies. In Döring, M. & Nerlich, B. (Eds.) *The Socio-Cultural Impact of Foot and Mouth Disease in the UK in 2001: Experiences and Analyses*. Manchester University Press.

Hagar, C. & Haythornthwaite C. (2005). Crisis, Farming & Community. *The Journal of Community Informatics*, 1(3), pp. 41-52.

Halfpenny, P. and Procter, R. 2009. Special Issue on e-Social Science. *Social Science Computing Review* 27, 4 (Nov. 2009), pp. 459-466.

- Hiltz, R. & Turoff, M. (1993). *The Network Nation: Human Communication via Computer*. MIT Press.
- Hine, C. (2000). *Virtual Ethnography*. London: Sage Publications.
- Hine, C. (Ed.) (2005). *Virtual Methods: Issues in Social Research on the Internet*. Berg,
- Hutchins, E (1995). *Cognition in the Wild*. MIT Press.
- Intel Research. (2006). Place Lab. <<http://www.placelab.org/>>.
- ISCRAM (2009). Information Systems for Crisis Response and Management, <<http://www.iscram.org/>>.
- Jones, S. (Ed.) (1999). *Doing Internet Research: Critical Issues and Methods for Examining the Net*. Sage Publications.
- Jurafsky, D. & Martin, J. *Speech and Language Processing, 2ed*. Prentice-Hall, 2008.
- Kendra, J. & Wachtendorf, T. (2006). Community Innovation and Disasters. In Rodríguez, H., Quarantelli, E. & Dynes, R. (Eds.), *Handbook of Disaster Research*: 316–334. Springer.
- Landgren, J. (2006). Making Action Visible in Time-Critical Work. In *Proceedings of the ACM 2006 Conference on Human Factors in Computing Systems (CHI 2006)*, Montréal, pp. 201-210.
- Liu, S., Palen, L., Sutton, J., Hughes, A., Vieweg, S. (2008). In Search of the Bigger Picture: The Emergent Role of On-Line Photo-Sharing in Times of Disaster. In *Proc. ISCRAM 2008*.
- Liu, S.B. & Palen, L. (2010). The New Cartographers: Crisis Map Mashups and the Emergence of Neogeographic Practice. Special Issue on Mapping Hazards and Disasters *Cartography and Geographic Information Science (CaGIS) Journal*, 37 (1), pp 69-90.
- Mark, G. & Semaan, B. (2008). Resilience in Collaboration: Technology as a Resource for New Patterns of Action. *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work (CSCW 2008)*, Nov., San Diego, 137-146.
- Mark, G. & Semaan, B. (2009). Expanding a Country's Borders During War: The Internet War Diary. In *Proc. ACM Workshop on Intercultural Collaboration*, Stanford University, 3-12.
- Mark, G., Al-Ani, B. and Semaan, G. (2009). Resilience through technology adoption: Merging the old and the new in Iraq. *Proceedings of the ACM 2009 Conference on Human Factors in Computing Systems*, Boston, pp. 689-698.
- Meier, P. & Brodock, K. (2008). Crisis Mapping Kenya's Election Violence: Comparing Mainstream News, Citizen Journalism and Ushahidi. Harvard Humanitarian Initiative, Harvard University: Boston). <http://irevolution.wordpress.com/2008/10/23/mapping-kenyas-election-violence>
- Mendonça, D. & Wallace, W. (2004). Studying Organizationally-Situated Improvisation in Response to Extreme Events. *International Journal of Mass Emergencies and Disasters*, 22(2): 5-29.
- Mileti, D., Drabek, T. & Haas, J. (1975). Human Systems in Extreme Environments. Boulder, CO: Institute of Behavioral Science, Program on Environment and Behavior.
- Miller, D. & Slater, D. (2000). *The Internet: An Ethnographic Approach*. New York: Berg.
- Mui, L. & Mohtashemi, M. (2002). A Computational Model of Trust and Reputation. In *Proceedings of the 35th Hawaii International Conference on System Sciences (HICSS 2002)*, vol. 7, p. 188.
- O'Keefe, E. (June 4, 2009). FEMA Encourages Public Participation. *The Washington Post*, <http://www.washingtonpost.com/wp-dyn/content/article/2009/06/03/AR2009060303404.html>
- Ohm, P., Sicker, D. & Grunwald, D. (2007). Issues Involving Network Monitoring. *Internet Measurement Conference*.
- Palen (2009). Keeping Up with the Future of Human-Computer Interaction: The Many Masters of Scholarship, Social Action Work, and Rapid Technological Change. *Human Computer Interaction Consortium 2009*, February 4-7, Winter Park, CO.
- Palen, L. & Liu, S. (2007). Citizen Communications in Crisis: Anticipating a Future of ICT-supported Public Participation. *Proceedings of the ACM 2007 Conference on Human Factors in Computing Systems (CHI 2007)*, San Jose, CA., pp. 727-736.
- Palen, L. & Vieweg, S. (2008). Emergent, Widescale Online Interaction in Unexpected Events: Assistance, Alliance and Retreat. *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work (CSCW 2008)*, Nov., San Diego, pp. 117-126.
- Palen, L., Hiltz, R. & Liu, S. (2007). Online Forums Supporting Grassroots Participation in Emergency Preparedness and Response. *Communications of the ACM*, 50(3): 54–58.
- Palen, L., Vieweg, S., Liu, S. & Hughes, A. (2009). Crisis in a Networked World: Features of Computer-Mediated Communication in the April 16, 2007 Virginia Tech Event. *Social Science Computing Review*, Sage, 27 (4), pp. 467-480.
- Powell, J. (1954). *An Introduction to the Natural History of Disaster*. U. Maryland: Disaster Research Project.
- Project EPIC (2009). *Project EPIC: Empowering the Public with Information in Crisis*, University of Colorado, Boulder and University of California, Irvine. <<http://epic.cs.colorado.edu/>>.
- Qu, Y., Wu, P. & Wang, X. (2009). Online Community Response to Major Disaster: A Case Study of Tianya

- Forum in the 2008 China Earthquake. In *Proc. 42nd Hawaii Int'l Conf. on System Sciences*.
- Reiter, M. & Rubin, A. (1998). Crowds: Anonymity for Web Transactions. In *ACM Transactions on Information and System Security*, 1(1), pp. 66-92.
- Schneider, S. & Foot, K. (2002). The Web After September 11. In Rainie, L., Schneider, S. & Foot, K. (Eds.) *One Year Later, September 11 and the Internet*, Pew Internet & American Life Project Report.
- Shklovski, I., Burke, M., Kiesler, S. & Kraut, R. (2008). Use of Communication Technologies in Hurricane Katrina Aftermath. Position paper for the *HCI for Emergencies Workshop at CHI 2008*.
- Shklovski, I., Palen, L. & Sutton J. (2008). Finding Community Through Information and Communication Technology in Disaster Events. In *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work (CSCW 2008)*, November, San Diego, pp. 127-136.
- Shneiderman, B. & Preece, J. (2007). 911.gov. *Science*, 315(5814): 944.
- Sicker, D. (2003). Privacy & Anonymity through Role Based Authorization, *31st Telecommunications Policy Research Conference*.
- Sicker, D. & McCoy, D. (2007). A Mechanism for Detecting and Responding to Misbehaving Nodes in Wireless Networks. *SDR Workshop, IEEE SECON*, pp. 678-684.
- Sicker, D., Ohm, P. & Grunwald, D. (2007). Legal Issues Surrounding Monitoring During Network Research. In *Proceedings of the 7th ACM SIGCOMM Conf. on Internet Measurement*, pp. 141-148.
- Sorenson, J. & Sorenson, B. (2006). Community Processes: Warning and Evacuation. In Rodríguez, H., Quarantelli, E. & Dynes, R. (Eds.) *Handbook of Disaster Research*, Springer, New York, pp. 183-199.
- Stallings, R. A. & Quarantelli, E. L. (1985). Emergent Citizen Groups and Emergency Management. *Public Administration Review* 45: 93-100.
- Stallings, R.A. (Ed.) (2002). *Methods of Disaster Research. International Research Committee on Disasters* (Xlibris Corporation).
- Starbird, C, L. Palen, A. Hughes & S. Vieweg (2010). Chatter on *The Red*: What Hazards Threat Reveals about the Social Life of Microblogged Information. *Proceedings of the ACM 2010 Conference on Computer Supported Cooperative Work (CSCW 2010)*, pp. 241-250.
- Starbird, C. & Palen, L (2010). Pass it On? Retweeting in Mass Emergencies. To appear in the 2010 Proceedings of *Conference on Information Systems on Crisis Response and Management (ISCRAM 2010)*.
- Sutton, J., Palen, L. & Shklovski, I. (2008). Backchannels on the Front Lines: Emergent Use of Social Media in the 2007 Southern California Wildfires. In the 2008 Proceedings of *Conference on Information Systems on Crisis Response and Management (ISCRAM 2008)*.
- Tierney, K., Lindell, M. & Perry, R. (2001). *Facing the Unexpected: Disaster Preparedness and Response in the United States*. Joseph Henry Press/National Academy Press.
- Torrey, C., Burke, M., Lee, M., Dey, A., Fussell, S. & Kiesler, S. (2007). Connected Giving: Ordinary People Coordinating Disaster Relief on the Internet. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS 2007)*, p. 179a.
- Vieweg, S, Hughes, A, Starbird, C, Palen, L (2010). A Comparison of Microblogging Behavior in Two Natural Hazards Events: What Twitter May Contribute to Situational Awareness, *Proceedings of the ACM 2010 Conference on Human Factors in Computing Systems (CHI 2010)*, Atlanta, GA, pp. TBA.
- Vieweg, S., Palen, L., Liu, S., Hughes, A. & Sutton, J. (2008). Collective Intelligence in Disaster: An Examination of the Phenomenon in the Aftermath of the 2007 Virginia Tech Shootings. 2008 Proceedings of *Conference on Information Systems on Crisis Response and Management (ISCRAM 2008)*
- Wachtendorf, T. & Kendra, J. (2005). Improvising Disaster in the City of Jazz: Organizational Response to Hurricane Katrina. *Understanding Katrina: Perspectives from the Social Sciences*, <understandingkatrina.ssrc.org>.
- Whitby, A., Jøsang, A. & Indulska, J. (2004). Filtering out Unfair Ratings in Bayesian Reputation Systems. In *Proc. 7th Int'l. Workshop on Trust in Agent Societies*.
- Wu, P. Qu, Y. & Preece, J. (2008). Why an Emergency Alert System isn't Adopted: The Impact of Socio-Technical Context. In *Proceedings of the 22nd Human Computer Interaction Conference*. Liverpool, UK.
- Wu, P., Qu, Y., Preece, J., Fleischmann, K., Golbeck, J., Jaeger, P. & Shneiderman, B. (2008). Community Response Grid for a University Campus: Design Requirements and Implications. 2008 Proceedings of *Conference on Information Systems on Crisis Response and Management (ISCRAM 2008)*.