Social media use by government: From the routine to the critical

Andrea L. Kavanaugh [a,1], Edward A. Fox [a,2], Steven D. Sheetz [b,3], Seungwon Yang [a,*], Lin Tzy Li [d,e,4,5,6], Donald J. Shoemaker [c,7], Apostol Natsev [f,8], Lexing Xie [g,8,9]

[a] Department of Computer Science, Virginia Tech, VA 24061, USA
[b] Department of Accounting and Information Systems, Virginia Tech, VA 24061, USA
[c] Department of Sociology, Virginia Tech, VA 24061, USA
[d] Institute of Computing, University of Campinas, Campinas, SP 13083–852, Brazil
[e] Telecommunications Res. and Dev. Center, CPqD Foundation, Campinas, SP 13086–902, Brazil
[f] Google, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA
[g] Australian National University, Canberra, Australia

ARTICLE INFO

ABSTRACT

Social media and online services with user-generated content (e.g., Twitter, Facebook, Flickr, YouTube) have made a staggering amount of information (and misinformation) available. Government officials seek to leverage these resources to improve services and communication with citizens. Significant potential exists to identify issues in real time, so emergency managers can monitor and respond to issues concerning public safety. Yet, the sheer volume of social data streams generates substantial noise that must be filtered in order to detect meaningful patterns and trends. Important events can then be identified as spikes in activity, while event meaning and consequences can be deciphered by tracking changes in content and public sentiment. This paper presents findings from a exploratory study we conducted between June and December 2010 with government officials in Arlington, VA (and the greater National Capitol Region around Washington, D.C.), with the broad goal of understanding social media use by government officials as well as community organizations, businesses, and the public at large. A key objective was also to understand social media use specifically for managing crisis situations from the routine (e.g., traffic, weather crises) to the critical (e.g., earthquakes, floods).

© 2012 Elsevier Inc. All rights reserved.

* Corresponding author at: Department of Computer Science, 114 McBryde Hall, Mail Code 0106, Virginia Tech, Blacksburg VA 24061, USA.
E-mail addresses: kavan@vt.edu (A.L. Kavanaugh), fox@vt.edu (E.A. Fox), sheetz@vt.edu (S.D. Sheetz), seungwon@vt.edu (S. Yang), lintzyli@ic.unicamp.br, lintzyli@cpqd.com.br (L.T. Li), shoemaker@vt.edu (D.J. Shoemaker), natsev@google.com (A. Natsev), lexing.xie@anu.edu.au (L. Xie).
[1] Department of Computer Science, 2202 Kraft Drive Mail Code 0902, Virginia Tech, Blacksburg VA 24060. Tel.: +1 540 231 1806; fax: +1 540 231 9218.
[2] Department of Computer Science, 114 McBryde Hall, Mail Code 0106, Virginia Tech, Blacksburg VA 24061. Tel.: +1 540 231 5113.
[3] Department of Accounting and Information Systems, Pamplin Hall 3080 Mail Code 0101, Virginia Tech, Blacksburg, VA 24061. Tel.: +1 540 231 6096; fax: +1 540 231 2511.
[4] This author was with Department of Computer Science, Virginia Tech, VA 24061 when this work was performed.
[5] Institute of Computing, University of Campinas, Av. Albert Einstein, 1251 – Campinas/SP – Brazil, 13083–852. Tel.: +55 19 3521 5887.
[6] CPqD Telecommunications R&D Center, Rodovia Campinas – Mogi-Mirim (SP-340) – km 118,5 – Campinas/SP – Brazil, 13086–902. Tel.: +55 19 3705 7136.
[7] Department of Sociology, 512 McBryde Hall, Virginia Tech, Blacksburg, VA 24061–0137. Tel.: +1 540 231 8971.
[8] This author was with IBM Watson Research Center, Hawthorne, NY 10532 when this work was performed.
[9] Current address: Bldg 108 Australian National University, Acton, ACT 0200, Australia.

1. Introduction

Citizens are increasingly relying on social media for communication with their family, friends, colleagues, businesses, and the government. The capabilities to facilitate interpersonal and group interaction provide new and unique opportunities for community leaders, elected officials, and government service providers to inform, and be informed by, the citizenry (Golbeck, Grimes, & Rogers, 2010). Twitter, Facebook, Flickr, YouTube, and other services with user-generated content have made a staggering amount of information available online. In 2010, during the period of our study, Twitter was generating an estimated 55 million tweets a day ("Twitter blog: Measuring tweets, 2010"), Flickr was amassing more than 6000 photos each minute ("Flickr, 2010"; Smarter Transportation: 10 Social Media Tools to Navigate Your City, 2010; "YouTube Blog, 2011"), YouTube was accumulating over 60 hours of video per minute ("YouTube blog. Holy Nyans! 60 h per minute and 4 billion views a day on YouTube"), taking up more than 10% of all internet traffic, and Facebook had more than 400 million active users; by early 2012 the number of Facebook users had more than doubled to 845 million (Protalinski, 2012), making it the most visited site on the Internet in the US. All this information and deep reach are readily available for government officials to tap into and leverage for improved

services. However, the treasure trove of information comes with substantial noise that must be filtered to make this information useful and reliable.

Government officials seek to leverage these resources to improve services and communication with citizens, especially segments of the population that previously were difficult to reach and underrepresented (Bertot, Jaeger, & Hansen, 2012). Yet, the sheer volume of social data streams generates substantial noise that must be filtered to be useful. The desire and potential exist for identifying and responding to issues in real time for more effective emergency management as well as improved public safety and general quality of life. For example, critical events of interest (e.g., earthquake, flash mob gatherings, protests, etc.) can be identified as spikes in the social media volume. Issues of concern for public safety or general quality of life (e.g., traffic, air quality) can be discovered, monitored, and mitigated by analyzing social media streams to detect meaningful patterns and trends (Fig. 1).

Similarly, monitoring these patterns and themes over time could provide officials with insights into the perceptions and mood of the community that cannot be collected through traditional methods (e.g., phone or mail surveys) due to a variety of reasons, including the prohibitive cost and limited reach of traditional methods as well as the limited window of opportunity for influencing or mitigating events as they evolve. Perhaps most importantly for emergency management, no traditional method can provide insight in real time. Surveys require substantial time and effort prior to data collection, during the collection process, and for analyses of the results, which often take months to complete. Secondly, substantive costs are associated with survey activities, making them especially difficult in light of reduced and shrinking budgets of governments at all levels. Finally, once completed a survey captures perceptions at a single point in time. Although it is possible to use surveys at intervals to monitor progress, it is not a common practice, substantially increases costs, and often does not reach important segments of the citizenry.

Data mining of diverse real-time feeds of social streams related to real-world events is needed to enable officials to make sense of the vast amount of information generated. In so doing, government should be able to act more effectively on matters both routine (e.g., ongoing issues of public concern) and critical (e.g., major weather or traffic disruption, public safety or rapid response). Using social media, we can answer questions that are not normally addressed by the gather-and-report style of journalism involving traditional sources, such as: When and where are events of importance currently happening? What are the different views of a given event? Who are the influential users in an online or local community? Yet, to use these resources effectively, we first need to address a series of questions including: Which social media should government use to communicate most effectively with a diverse public? How should messages be formed and framed across social media to be effective? To what extent can messages in social networks be used to explain how influential messages form and spread? Is civic information, disseminated through social media as opposed to through the Web or email, more likely to reach some traditionally underrepresented groups, such as those with lower socio-economic status (SES) or younger voters? What role do social media play in the general mix of information sources for citizens to communicate about civic life, with each other and with government? Do social media affect civic participation and if so, for whom and what kinds of civic participation?

We seek to leverage technology to help government manage information and facilitate interaction in meaningful ways in order to achieve broader public participation than is possible through normal channels (e.g., public commenting at county board meetings). Deep analysis of social media streams can also provide access to segments of the community that have not participated in traditional ways.

This exploratory study was part of a larger investigation funded by NSF (IIS-0916733) to build a Crisis, Tragedy, and Recovery Network (CTRnet) ("CTRnet: Crisis, Tragedy, & Recovery, 2010"). In collaboration with Arlington Virginia County government, we conducted a six-month exploratory study of how social media were being used by local citizens, community organizations and government, and how data analysis could be applied in Arlington and environs to improve services and communication with citizens. Our primary research objectives were to investigate the use and impact of social media and to identify and develop methods to effectively meet a variety of local government and community needs.
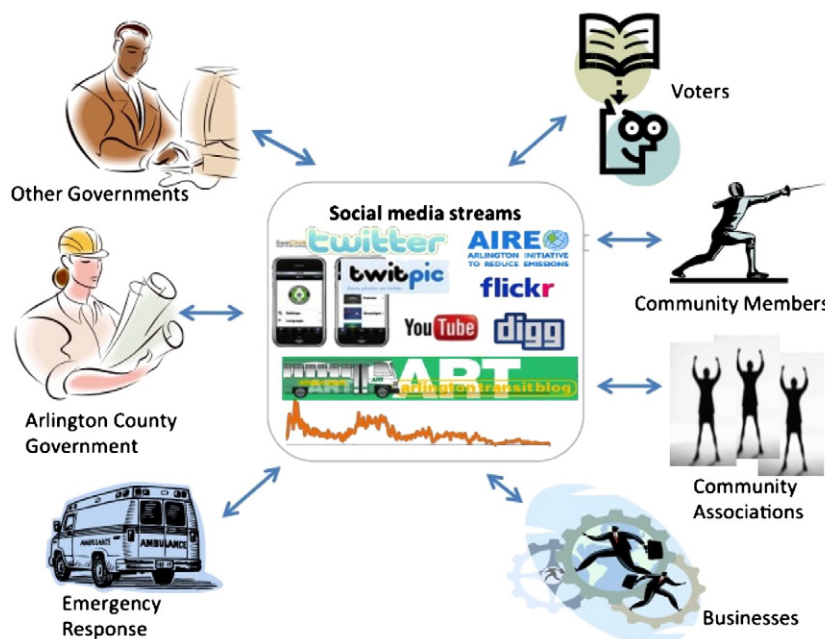

Fig. 1. Social media streams to improve services and communication with citizens.

Specifically, we have begun to:

1) Leverage and further refine tools for collecting and correlating large amounts of public social media data relevant to Arlington County, VA and environs,

2) Archive and curate collected social media data over a period of time into a digital library, including social media for crisis conditions, and

3) Identify, research and implement applications of multimedia analytics and text mining for government services and communication.

To address these goals we crawled, collected, aggregated, and archived relevant social media data; we conducted exploratory focus group interviews with key stakeholders in government and community organizations, and developed tools to analyze and render data more usable and meaningful for local organizations, governments and citizens.

Our target information sources included official Arlington County Facebook pages, Twitter feeds ("Arlington County Blog Central, 2010"; "Arlington County Facebook Profile, 2010"; "Arlington County Flickr account, 2010"; "Arlington County News on Twitter, 2010"; "Arlington County & VA Official Site, 2010"; "See-Click-Fix in Arlington County, 2010"), blogs, news, community forums, and relevant postings by the public on social media sites such as Twitter, Facebook, YouTube, and Flickr. Applications of such analyses could include monitoring public opinion before and after large public events, monitoring planned or unplanned activities, identifying and categorizing important community issues over time and location, enhancing community recovery in response to crises or tragedies, and tracking the development of long-running themes in civil life.

While many government agencies have recognized social media as an important information source and outlet, there has yet to be a comprehensive account about the needs and methods for social media use. Recent case studies, such as published by Queensland Police Service ("Queensland Police Service, 2011") outlined the experiences and best practices for engaging and informing citizens during a historic flood. To the best of our knowledge, this study is the first to survey across a wide range of government agencies and community organizations, supported by data analysis of existing online interactions.

The CCSR is a partnership among Virginia Tech, IBM, and Arlington County. Based on interests and needs demonstrated in a CCSR workshop with officials from Arlington County and the National Capital Region (NCR) (the area around Washington, D.C.), we planned the exploratory study in collaboration with IBM and Arlington County government to explore social media applications that might improve community resilience in times of crises, as well as provide timely and complementary open sources of information for facilitating city, county, and community services. Further, we explored social media applications that might help agencies make sense of a deluge of information by providing meaningful consumable insights.

2. Social Media and Government

Social media are internet-based applications designed to facilitate social interaction and for using, developing and diffusing information through society. Social media build on many of the same concepts and technologies of Web 2.0, most basically, the creation and exchange of user generated content (O'Reilly, 2007). There is much overlap between the two concepts and technologies in terms of examples, including blogs, wikis, and recommender systems; websites to share videos, music, pictures and podcasts; and social networking sites such as Facebook and MySpace. Broadly, Web 2.0 and social media are considered social software, i.e., software that enables people to rendezvous, connect, or collaborate through computer-mediated communication (Boyd & Ellison, 2007; Lampe, Ellison, & Steinfield, 2006). This type of software has existed for years in the form of online

bulletin board systems listservs, forums, and newsgroups. More recently, however, blogs (Tepper, 2003) and microblogs (e.g., Twitter), RSS feeds, tagging systems (Furnas et al., 2006), and collaborative filters have made social software easy to use and highly scalable leading to greater adoption and use.

2.1. Social Media Use by Citizens

Social media have changed the way many Americans get information about what's going on in their communities, and national and global current events. They provide new ways for citizens to share information and to interact with each other and with elected officials and government agencies. A national study conducted by Pew Internet & American Life in 2010 finds that almost a third (31%) of all online adults in the USA used social tools such as blogs, social networking sites, and online video as well as email and text alerts to keep informed about government activities (Smith, 2010).

Social media seem to have particular appeal for groups that have historically lagged in their use of other online government offerings — in particular, minority Americans (Smith, 2010). Latinos and African Americans are just as likely as whites to use these tools to keep up with government, and are much more likely to agree that government outreach using these channels makes government more accessible and helps people be more informed about what government agencies are doing. Findings from the 2010 Pew study also show that 40% of adult Internet users have gone online for raw data about government spending and activities. This includes anyone who has done at least one of the following: looked online to see how federal stimulus money is being spent (23% of internet users); read or downloaded the text of legislation (22%); visited a site such as data.gov that provides access to government data (16%); or looked online to see who is contributing to the campaigns of their elected officials (14%).

In a 2009 online convenience sample survey conducted in the US by the American Red Cross ("American Redcross: Social Media in Disasters and Emergencies"), 75% of respondents reported they would use social media in crisis and civic-related situations (e.g., traffic jam, car crash, potential crime, or downed power lines). Nearly half of respondents reported that they would use social media to let others know they were safe in an emergency; 86% report they would use Facebook; 28% would use Twitter, and 11% would use a blog. Solutions that are (already) provided by the industry for public safety include call processing products and notification systems. For example, Plant CML offers call processing software that is used by 2/3 of all 911 centers in North America. They also provide notification systems, computer-aided design & mapping, data management and analysis, information management, and land mobile radio. These systems, however, are mostly based on phone communications and are not using the power of social media.

Large public gathering events, such as parades or demonstrations, are examples of conditions of social convergence, that is, high-intensity events with large population density and heightened security needs. Before the event it is beneficial to monitor online discussions on national and global sources, such as YouTube and Twitter, as well as local sources, such as Arlington blog central ("Arlington County Blog Central, 2010"), local Facebook pages, YouTube and Twitter posts ("Arlington County Facebook Profile;" "Arlington County News on Twitter;" "Gasbuddy: Find local gas prices, 2010"), or Foursquare "check-ins" (or similar location-aware mobile media applications). This monitoring helps community leaders and the public stay informed about the various perspectives, sentiments, feedback, and insights around an event or an issue of interest. Afterwards, if a security event has emerged (e.g., violence or vandalism), sometimes evidence will be posted on photo and video sites, which can help local officials identify and track suspects as an event progresses. In epidemic propagation and prevention, on the other hand, the focus of information management is on early spotting of cases and managing

public input, contributions, and feedback around issues like quarantine, vaccination, and distribution of sanitary advice (e.g., swine flu).

Research on the use of Twitter in crises has a short history, as Twitter was established only in 2006. A form of micro-blogging, Twitter is a free, short messaging service with some social networking features. Some of the most relevant work to ours has been done by Palen, Hughes, and colleagues (Hughes & Palen, 2009; Hughes, Palen, Sutton, Liu, & Vieweg, 2008) and by Zuckerman on the Moldovan election protests in Africa (Zuckerman, 2009). These studies specifically focus on the use of Twitter during disasters and conditions of social convergence, such as mass political demonstrations, rallies or riots. Hughes et al. (Hughes & Palen, 2009) report that Twitter use under duress and in crisis conditions of the two hurricane episodes of Ike and Gustav in 2008 is distinct from routine general Twitter communication behavior in two ways: 1) fewer tweets are sent as replies to other tweets; and 2) fewer URLs are included in the tweet posts. They surmise that this is because in a crisis, people need to broadcast information as widely as possible to as many people as possible at once (i.e., no need to reply to a specific individual) and people are less likely to go to a website for additional information during an emergency. As an emergency is unfolding, Twitterers may broadcast more up to date and timely information (and sometimes misinformation) than government organizations and mainstream media that take the time to double-check the accuracy of their public information, especially during a crisis.

2.2. Social Media Use by Government

Twitter and other social sources have been effective in early event spotting (Opsahl, 2010; Sakaki, Okazaki, & Matsuo, 2010), the response time of which can be even faster than official sources (e.g., earthquake reporting). Such monitoring strategies also can be used for epidemic spotting and trending, where monitoring should be both distributed and spanning a longer period of time, such as the first case in each school district, resurgence of disease cases, and long-range planning for local management. In the case of continuous monitoring, social media can help measure the effectiveness of control measures and propaganda, e.g., if the public is embracing the vaccine distribution scheme, complaining about it, or helping authorities stay better informed about gaps or deficiencies in its administration.

We have been studying social media use and impact as part of an ongoing longitudinal investigation of Internet use and impact in Blacksburg, Virginia and environs since the early 1990s ("Social Media Sells, 2010"). Blacksburg is home to the main campus of Virginia Tech (which also has a small campus in northern Virginia near Arlington) and is home to the community computer network known as the Blacksburg Electronic Village (BEV). Blacksburg town government has won several awards for its rich mix of media to inform and communicate with citizens, including Twitter and Facebook since January 2009 as an additional channel for 'Blacksburg Alerts' available by email or text message. The Communications Specialist in town government monitors Twitter (using TweetDeck) for relevant posts that would benefit from a reply (e.g., "the town does not have control over the old middle school in Blacksburg, that is the County's jurisdiction") or should be brought to the attention of town council as a citizen suggestion ("it would help to have a cross walk painted at this intersection; it's very busy").

While this was not the case for the town of Blacksburg government, in the National Capitol Region focus group participants noted that the public relations person for various government agencies was typically not familiar with nor comfortable with social media. This limitation makes it especially difficult for the public relations office to manage this channel of communication with the public.

From our preliminary study of social media use in Blacksburg, we found that most often the person posting tweets or managing an organization's Facebook page was not from the organization's leadership. Instead, a college student or other young adult was often working in tandem on behalf of the organization to post announcements,

updates, or other information. Some other US communities, including our project partners at Arlington County, Virginia, are experimenting with monitoring Twitter and Facebook using a Web tools like TweetDeck and Hootsuite, in order to monitor social media communications and potentially to reduce workload and enhance responses at 911 centers (Opsahl, 2010).

3. Study Methods

We collected and analyzed area-specific social media (social media) sources, and conducted focus group interviews with 25 county officials (specifically, personnel from emergency management services, the police department, and volunteer leadership office), including a questionnaire about their social media use and community involvement. We were able to recruit 25 participants and organized them into three separate focus group sessions (lasting two hours each) held in November and December 2010 in Arlington. At the outset of each of the interview sessions, we asked participants to complete an online questionnaire. The questionnaire asked them about their use of social media and their involvement in the local community.

The focus group sessions proceeded in two main stages. They began with participants engaged in electronic brainstorming to generate a number of ideas quickly, followed by a process whereby they identified categories that grouped the ideas by similarity.

Using individual computers with group support software that we developed, the focus group participants anonymously generated and entered ideas, beliefs, issues, or concepts, in the form of short sentences or phrases that they felt were important to the situation. We provided them with a set of framing questions we developed to cue participants to begin entering ideas. Fig. 2 shows the framing questions we used in the focus groups. The ideas participants generated were shared with other participants as they were generated, allowing ideas generated by one person to be expanded by others or to cue others to generate related ideas. Participants then worked together with the facilitator to create and name the units or categories that organized their ideas by similarity.

We collected social media in the form of official posts and public comment data from the Arlington County Facebook page, Twitter feeds from local civic organizations, YouTube videos, and crawls and searches of local web pages. We used different Twitter analytical tools, such as '140 kit' (http://www.140kit.com) and the Archivist (http://archivist.visitmix.com) to collect tweets from 34 local organizations, including Arlington government, that were civic in nature (rather than commercial or residential).

We performed semantic analyses on the Twitter data to identify popular topics and to characterize followers by their profile data; we conducted simple frequency counts to calculate number of 'followers' and 'followers of followers' of a given organization. We used the visualization software 'wordle' (http://www.wordle.net) to represent the results of the Twitter analyses as tag clouds in order to be able to distill and make greater sense of large amounts of data more quickly and easily. For the YouTube video collections, we used a Perl script to search all YouTube videos for the tags or video with the title 'Arlington County' and represented the search results in a tag cloud indicating the most frequent tags in the image collection.

4. Results

Our findings from the exploratory study are based on the focus group interviews and participant questionnaires (N = 25), and the development of tools to analyze social media data we collected. The results fall into three main areas:

1) Local government uses social media without knowing its costs and benefits, or who their actual audience is, who in their organization should monitor communications, how and when they should be

- What are the missions and objectives of your organization?
- What are you trying to accomplish using social media?
  - Do you feel you are currently accomplishing this goal effectively with social media? (if yes, why?)
  - If not, what do you need [to know?– to do? --in order] to use social media more effectively?
- What concerns do you have about using social media?
- What difficulties do you have about using social media?
- What information would you like to have about how your organization uses social media?
- What information would you like to have about how social media is being used in your community?
- Is there anything else you would like to know about social media that would be helpful?

Fig. 2. Framing questions for focus group interviews.

responding, and what effect their social media communications have on the public;

2) New tools are needed to help government and citizens make sense of the overwhelming amount of data that is being generated, to model the flow of information, and to identify patterns over time; and

3) Digital libraries are needed to archive and curate generated content, especially for crisis and social convergence situations, but also for analyses that cover longer time frames.

4.1. Focus Group Questionnaire

Each of the 25 focus group participants completed an online-questionnaire at the outset of their focus group interview session. Of this sample, 15 (60%) were female and 10 (40%) were male. The majority (84%) was white, non-Hispanic. Sixty-four percent were married and 92% were employed on a full-time basis. It is reasonable to characterize the interviewees as civically active, as they reported being very active in their community and being well informed about local news and politics. Seventy-six percent reported that they kept up with local news daily.

Most respondents reported they have ideas for improving things in their community at least once a month (76%) and that they frequently get together with others who are well informed about local issues. Thirty-six percent reported that they worked to bring about change in their community on a daily basis. Almost half of the participants (48%) reported that they either posted comments online, posted pictures or video online, or blogged about a political or social issue in the past year.

The overwhelming majority (80%) of respondents reported having a profile on at least one type of social media website (social networking site, blog or Twitter, photo/video collections, place-based applications, or other). All of these profile-users maintained a profile on a social networking site, with many having profiles on multiple types of social media sites.

Respondents used social networking sites more frequently than other types of social media sites. Over half (56%) of respondents used social networking sites on a daily basis, and 76% used these sites at least once a week. Place-based applications were the least used type of site. Of the 5 individuals who used these applications, none used these sites more than once a month. Most respondents accessed these social media sites via personal computer (96%) and many used their cell-phones, too (68%).

Social media use was fairly well distributed across types of social media sites, with the exception of place-based application (social networking sites 56%, blog or microblog 44%, and photo/video collection 40%). All in all, 64% reported using social media sites to communicate with other members of their organization, with several respondents utilizing multiple types. Fifty-two percent reported using social media sites for such purposes at least once a week.

The respondents were generally satisfied (88%) with current emergency response efforts in their community. All respondents felt that the county government should contact citizens by way of phone call or text message during a crisis. Eighty-four percent felt that social networking sites also should be utilized for this purpose, and 72% felt that blogs or microblogs (e.g., Twitter, Tumblr) should be as well. Over half (56%) of respondents reported that they were at least somewhat likely to use one or more types of social media to contact family members during a crisis. However, only 24% were likely to report a crisis to local government agencies via social media. The majority of respondents reported that talking to others in person or by telephone was the most important source of local information.

4.2. Focus Groups: Information Factors

In the electronic brainstorming step of the focus groups, participants identified 23 categories of factors related to 1) the organization and 2) the information exchanged between the organization and community (Fig. 3). Information factors include issues related to the quality and quantity of information generated through social media. They also include the tone and types of communications in which government seeks to participate, including outreach, feedback, and two-way communication. Additional types of information that can be obtained from some social media channels, e.g., detecting the locale of emerging events, are of substantial interest for emergency management and policing functions. Finally, technology issues included the security of the technology used to provide social media and other new tools, and the need to meet legal obligations for saving public records. Further, substantial questions remained among participants regarding which social media should be utilized for diverse audiences and purposes.

Together the factors identified by the participants describe a broad range of interests and concerns of the Arlington County government in relation to their use of social media. Each of these categories also contains a set of ideas from the electronic brainstorming that further clarifies the intentions of the participants about the meaning of the categories.

4.3. Focus Groups: Organization Factors

The organization factors that focus group participants identified include policies, legal issues, costs and training (Fig. 3). The organization requires that polices be adopted to provide the environment needed for employees to achieve work objectives. Management buy-in is essential if benefits are to be realized and costs are to be controlled. To utilize social media effectively, employee activities and roles are institutionalized through Human Resources (HR) to clarify job descriptions and ensure related types of communication are managed effectively. There are also attempts to control information and to communicate the government's opinions and actions to the public.

❖ **Information Factors**
  ➢ Communications
    ▪ Community Outreach (emergency, crime/traffic alerts, 24/7 level of service, recruitment)
    ▪ Feedback (from community to organization, social trends, locale, fast spreading ideas)
    ▪ Population Reached (misses traditional/older population or can't afford technology)
    ▪ One Way vs. Two Way (pushing out vs. creating dialogue, effort/costs different)
    ▪ Tone (Government presents just the facts, not stories, not press release, listen then educate)
  ➢ Information
    ▪ Quality of Content (accuracy, facts of situation, un-vetted information, misinformation)
    ▪ Quantity of messages (how to be heard, from 1 to 10 to 1000s, overwhelming, loss of control)
    ▪ Personal Level (information overload, ability to write complete thoughts, nuances of face-to-face lost)
  ➢ Technology
    ▪ Security (network exposed to world)
    ▪ Technology and Equipment (cost of technology and maintenance, cost savings, training)
    ▪ Social Media (SM) Outlets (knowing audience/expertise, users expect transparency, so many outlets)
    ▪ Public Record/FOIA (are SM public record, tools needed to save, outdated polices)

❖ **Organization Factors**
  ➢ Policy
    ▪ Management Buy-In (unknown expectations, under valued, need to set culture)
    ▪ Control Issues (how much to control, what we can control, telling how/what to think/do)
    ▪ Human Resource (HR) Components (job descriptions, evaluation, expertise, dialogue, positive and negative)
    ▪ SM Communications Policy (what not to do/say, right people to make SOP, moving target)
    ▪ Professional Level (privacy concerns, devices owned by county, investigative purposes)
  ➢ Legal Issues
    ▪ Data Maintenance (FOIA data maintenance and related costs)
    ▪ Owing Vs Using Someone Else (official outlet versus imposter, use in investigations.
    ▪ Public Record/FOIA (are SM public record, tools needed to save, outdated polices)
  ➢ Costs
    ▪ Resource Issues (SM adds to previously full time job, other duties, limit 24/7 expectation)
    ▪ ROI/cost to value (how to measure value, who are we reaching, enough received messages)
  ➢ Training
    ▪ Education (tools to manage, learning from each other, train constituents where to go)
    ▪ Training (best practices for dividing duties, case studies, understanding management's concerns)
    ▪ Other (educate nonusers, establish boundaries)

Fig. 3. Simple taxonomy of categories identified by focus group participants.

Organizations seek to define the types of information to be shared and the manner of sharing. The participants perceive the substantive legal issues related to maintaining government transparency, often through the Freedom of Information Act (FOIA), as important in considering the use of social media. For example, should tweets by a government employee be part of the public record? What about tweets by a government employee that are related to their non-work life? The individual government employee should set up two different identities in Twitter, in order to separate professional and private roles and representations.

Costs are always important to organizations, and government budgets have been squeezed due to reduced receipts resulting from slowing economic activity and increased use of government services. Yet the participants perceive that the potential exists for achieving efficiencies using social media and the potential return on investments should be evaluated. Complicating this calculation is the value placed on reaching previously uninvolved constituents and the most interested participants. One of the costs of adopting social media is related to training the employees who will use them. In addition, the public must be educated to understand how the government will interact with them and what expectations for interaction are appropriate.

Some of the Arlington County focus group participants said that they need social media aggregation tools. In general, dashboard services that accept search keywords and phrases help monitor information from multiple social media, such as trackur (http://www.trackur.com/social-media-monitoring) and Netvibes (http://www.netvibes.com/). But these tools are designed to support businesses not government or citizens, so they are not optimal for civic needs. Having geo-mapping features would be very useful for the needs of cities and communities, which are not currently enabled in dashboard tools.

Some emerging applications allow citizens to contribute geo-tagged photos and video to a community database. For example, MIT's Mobile Media Experience Laboratory has developed a place-based application called Locast for this purpose (http://www.locast.mit.edu). The video analytic software IBM has developed will help to organize and cluster images of similar content or location. This would make it easier for users to find content of interest and to contribute to ongoing information exchange regarding a particular issue related to a specific place (e.g., building a new school).

Some focus group participants also indicated that recent or projected budget cuts could erode 15 years of community outreach; the County wanted to understand how to use technology to maintain and sustain established communications with citizens. The Arlington area homeowners and neighborhood/civic associations have been key links for government to achieve community outreach in the past, but not all neighborhoods have homeowners associations. Residential neighborhoods lacking associations are usually characterized as having lower socio-economic status (SES) as measured by education and income; with budget cuts it is even harder for government to sustain routine outreach and communication with these areas of the city. Social media may be particularly helpful for outreach to such households and neighborhoods, especially through neighborhood opinion leaders and cell phones.

Preliminary evidence from a national study by Pew Internet & American Life (Smith, Verba, Brady, & Schlozman, 2009) indicates that the use of social media for civic purposes is not as strongly correlated with education and income as the use of traditional internet (i.e., web browsing and email). This may be because opinion leaders (i.e., influential individuals) exist at all social strata, and they may convey information to members of their social circles not only face to face, but also by cell phone. Cell phone ownership permeates all social strata and exceeds computer ownership among lower SES groups.

The cell phone is essentially a pocket computer. For lower SES groups it is likely to be the only computer they are using. While we were not able to study cell phone use among lower SES groups in this exploratory study, we are investigating cell phone use for civic purposes among similar demographic groups in southwest Virginia in related research. We are investigating the possible use of cell phones to address information needs, and their connection to social media, especially text messaging and image sharing between government, local opinion leaders and lower SES populations.

4.4. Tools for Analyzing Social Media Data

In order to study the pattern of communication and the information communicated using social media, we collected publicly available data from Twitter. We identified 34 civic organizations, some of which are government agencies, in the National Capitol Region (NCR) that were tweeting; we collected and analyzed their tweets for 30 days between September and October 2010.

We analyzed the tweets as well as the biographical information posted as profiles of the organizations' followers using the Natural Language Toolkit, tag clouds, and graphs. Fig. 4 shows the number of followers for the 34 civic organizations in the NCR. In order to get a sense of who are the followers of these 34 civic organizations, we collected the publicly available biographical profile information that followers list on their own Twitter accounts.

For the 34 civic organizations that were tweeting during the September–October 2010 period, we see there were a total of about 31,000 'direct' followers (i.e., people who subscribe to the RSS feed that carries each organization's Twitter posts). What is interesting to note is that the 'direct' followers are themselves being 'followed' by other people — what we refer to as 'followers of followers.' The number of followers of followers for these same organizations is over 67 million (Fig. 5). By looking into the number of followers of followers, we see the great extensibility of the communication chain radiating out beyond the original tweet.

One of the 34 civic organizations, a local news and events group called Arlington Unwired (Arlington UW) is shown with an arrow in Fig. 4. It had only 471 followers on the date we captured these data (September 26, 2010). We can see from the analysis of the number of Arlington UW followers' followers (Fig. 5) there are over 8 million followers. This is not to say that a tweet from Arlington UW will go beyond the 471 direct followers; however, if there is a crisis in the Arlington area (such as a major catastrophe or extreme violence) it is very likely that the *indirect* followers will retweet (forward along the same Twitter post) regarding such a catastrophe to their own set of followers (i.e., over 8 million followers). In this way, a critical piece of information has the potential of being disseminated throughout a community far beyond the direct followers to a larger population of followers' followers.

It is also important to note that among the followers of Arlington UW is 'Barak Obama' — and the number of followers of 'Barak Obama' is over 5 million. Further analyses tell us which words are used most commonly in the Arlington UW followers' bios during this period. The predominance of various words (most common words appear larger in a tag cloud) provides a quick overview of what is being said or characterized (in the case of followers' bios). Fig. 6 shows in a tag cloud the predominant profile descriptors given by the followers of Arlington UW.

This kind of profile analysis and visualization provides a quick overview of the type of individuals and their interests who are following a given organization. We analyzed and visualized in a tag cloud the 20 recent tweets of the followers of Arlington Unwired (UW) at the time of the data capture (September 26, 2010) shown in Fig. 7.

By looking at the recent tweets of Arlington UW followers, we see a kind of 'mood' and 'buzz' among users. The many references to time, e.g., today, tomorrow, tonight, weekend, reveal a focus on things happening around the period that the tag cloud is generated. However, the specific events presumably included in the tweets do not occur often enough to be included in the tag cloud. The large 'RT' stands for re-tweet, meaning that this is the most common term appearing in the Twitter posts for these users. The organization knows from this analysis that many of their posts are going well beyond their immediate (direct) followers.

Fig. 8 shows approximately 40 frequent words from the follower profiles of CarFreeDiet, which is one of Arlington's commuter services that promote healthy and environment-aware life-styles without (or less use of personal) cars. The organization's Twitter profile says, "Arlington's Car-Free Diet is the easy, fun way to live a car-free lifestyle".

Fig. 9 shows the most frequent words from the 20 'latest' tweets (at the time of data capture, September 2010) of the followers of
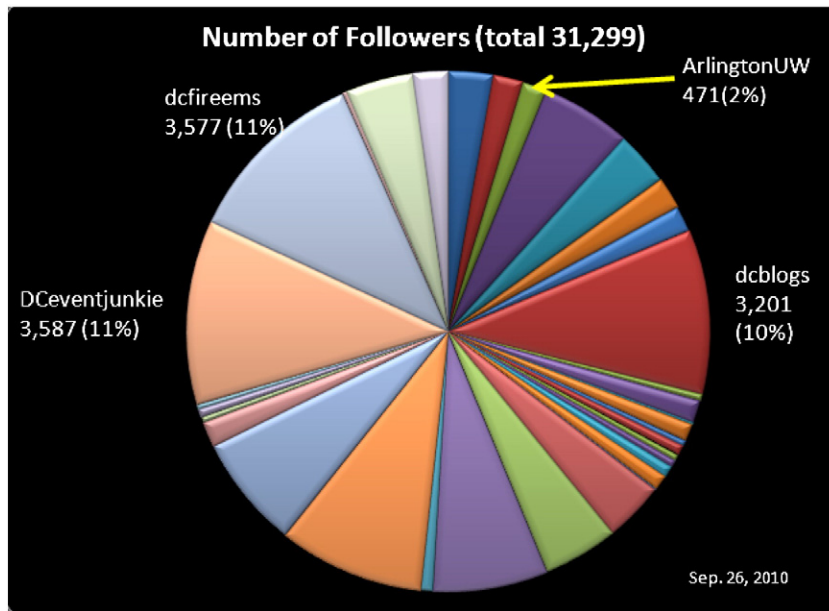
Fig. 4. Number of follwers for 34 NCR civic organizations.

CarFreeDiet; again the symbol for re-tweet (RT) was the most prominent. Therefore, the organization knows their information is actively shared and disseminated throughout the community because their Twitter followers are re-tweeting the organization's tweets.

In order to have some insights regarding the geographic locations of the followers, we selected followers from the top 5 civic organizations in Arlington area, which had the most number of followers. Then we visualized their location information on a map.

Fig. 10(A) shows the macro-view of the location distribution. For example, it shows that the followers are distributed even in other continents such as Africa and Asia. When the icons in (a) are pressed, the visualization zooms-in and shows detail locations (Fig. 10(B)).

The purpose of these analytical and visualization tools, as noted earlier, is to allow government and citizens to see quickly and easily the big picture of the information and communication flows that interest them.

4.5. Analysis of Facebook Comments

Arlington County government has maintained a Facebook page since early 2010 (http://www.facebook.com/ArlingtonVA). The page had roughly 4500 fans at the end of September 2010. We analyzed a two-month period (August 1–September 30, 2010) of posts by the County and responses (comments) from the public by conducting a simple content analysis by topic. There were a total of 112 posts; the top 10 most frequent topics are shown in Fig. 11.

The most common posts by the County on the Facebook page were about traffic (e.g., conditions, closures, metro outages), followed by public service announcements (PSA). News (updates, and other County announcements) and weather related posts (National Weather Service and Arlington Weather Service advisories) were followed by various events (biking paths, walking, music or film) in terms of frequency of posts. There were only a few posts related to education (Arlington

Fig. 5. Followers of organizations' followers.

Fig. 6. Profile biographies of Arlington UW followers.

Fig. 8. Profile biographies of CarFreeDiet.

County School District) and library services (e.g., closures, speakers, special activities) during this two-month period.

There were a total of 824 public comments to the County posts during this two-month period. Half of the comments pertained to about a fifth (19%) of the County posts (the top 21 posts by the County). Fig. 12 shows the distribution of the bulk of the comments on the same top 10 County posts seen in Fig. 11.

The comments are predominantly related to traffic and miscellaneous events (that is, events that do not fall into the other 'event' categories shown, such as food, exercise, music, and film). Exercise events (biking, walking) and news announcements generated the next most frequent number of comments from the public. Almost all the comments were highly consistent with the social media policy of the County (e.g., no profanity or off topic comments) and were overwhelmingly positive in tone, including many "Likes".

Lastly, we collected videos in YouTube pertaining to Arlington, Virginia and conducted a tag analysis of the video collection using image software developed by IBM. We performed a search using a Perl script and the phrase 'Arlington County;' this produced about 1800 videos from YouTube. We generated two types of tag clouds using video titles and tags (see Fig. 13).

As noted earlier, a tag cloud as visualization quickly and easily represents the frequency with which different terms appear in a search thereby providing a snapshot of what is in a large dispersed collection. The more frequently a term appears in an image collection, the larger it appears in a tag cloud. The cloud visualization also provides an indication of the importance of various civic issues to members of the community. The recurring civic themes revealed in the video analysis can be further explicated in the six categories shown in Table 1.

A further clustering of video tags and video titles as shown in Table 1 allows government and other users to make sense more easily

of the interests and needs of the community as expressed in the YouTube collection at any given point.

5. Discussion and implications

The exploratory study was intended to advance technologies and systems for social media analysis relating to both routine day-to-day civil life and critical incidents or emergencies. The results begin to identify and address a combination of technical and social science challenges. On the technical side, these include:

1) Recognizing relevant information accurately and in a timely manner, especially from short content micro-blogging sites (e.g., Twitter); the limited information in a tweet (i.e., less than 140 characters) makes it difficult to identify its meaning and context which may lead to incorrect classification and misleading analysis of data;
2) Alerting government officials to the analyzed information from multiple social media sources; due to the massive volume of the social media data stream, it is a challenge to quickly analyze the collected information from different sources and to make a decision based on the analysis; and
3) Visualizing the current and past status of incoming information and the analysis of it; simple yet informative visualization design is essential in making-sense of the data presented. We support the sense-making process by incorporating interaction methods with visualization to deal with large amounts of data.

On the social science side, our exploratory study results build on social network analysis and social and political participation research on the use of social media. We also seek to contribute to crisis informatics research and an understanding of the use of social media in crisis situations, including more mundane crises, such as major weather or traffic problems, and in social convergence situations,

Fig. 7. Twenty recent tweets by followers of Arlington UW.

Fig. 9. Twenty recent tweets by followers of CarFreeDiet.

Fig. 10. Follower locations visualized on a map. (A) Zoom-out view; (B) Zoom-in view for more detailed location information.

such as crowds, rallies, and other large gatherings that are not unusual in the National Capitol Region (NCR).

We focus on Arlington and the NCR as our test case in order to analyze information, its use and impact related to local, state, national, and international events — since it has close connections to the US capitol.

Our social media data analyses are intended to help government and citizens of Arlington County and the NCR know how and where to get useful information and critical communication in the event of a crisis or social convergence condition. Our tools should help government and citizens monitor and make sense of the diversity of voices and information that enrich the quality of life in their communities. Tools we are developing will be available in open source for government and citizens to help them find information clustered by topic or place and to further contribute, discuss, and interact with each other.

By mining content and services covering multiple media types (i.e., text, audio, image, and video) we can develop tools to recognize events and alert government, citizens, and community groups to see quickly the 'big picture' through visualizations of social media activity and content and changes in both over time. The intent is to enable proactive responses, as routine problems or crises start to loom, as events unfold, as individuals and groups respond, and as plans (short or long-term) are made for improved services and communication. Such capabilities are relevant to a broad range of governments throughout the US and globally. Given the efficiency of communication provided by social media, coupled with the potential to reach many constituents quickly, governments should seek to understand and to leverage these increasingly popular communication channels.

Governments, local organizations and citizens will continue to use a combination of traditional communication methods (e.g., newspaper,

radio, television, magazine, telephone) and emerging tools, smart phones and social media. Governments know they have diverse audiences with different needs and preferences. Social media are just another set of communication channels to get word out and serve the interests of different (mostly younger) citizens. Citizens will continue to use different media to get and share information, not only with each other, but with government. There is a kind of ecology of tools and devices that interplay to meet various needs for multiple purposes and types of users.

That said, however, the interplay of traditional and emerging media may become quite blended over time. Until then, there are persistent costs (e.g., policies, legal concerns, and other issues) and benefits about which governments especially seek guidance. Typically, a public relations person handles communication between government and the public. But many public relations managers are not comfortable with social media. It is possible that a public relations manager could focus on the traditional communication media such as, newsletters, press releases, and phone interviews with local TV and radio. To manage communication with a more diverse public, however, they need someone who uses social media. So, either they have to re-train current public relations managers, or they have the added cost of adding another person to manage public relations activities that involve social media.

There will continue to be legal issues concerning the interaction between traditional communication methods and social media. These include managing different ways that the public can report a problem to authorities. Government is legally on notice of a pothole or downed power line whether someone phones it in or tweets it. How long do governments need to store tweets as part of the public record?

Increasingly, traditional communication technology (newspaper, radio, television, magazine, telephone) are digital and accessible

Fig. 11. Facebook topics Arlington County.

Fig. 12. Public comments by Arlington Facebook topic.

Fig. 13. Tag cloud of Arlington YouTube videos.

Table 1
Tag cloud categories for Arlington videos.

Law enforcement	Police, cops, officer, courthouse, robbery, accident, ACPD, surveillance
Transportation	Metro, street, boulevard, highway accident, parking, transit
Social issues	Environment, diversity, community, city, neighborhood, accountability
Economic development	Growth, sustainability, development, bank, private, local
Political	Government, elections, agencies, department
Communication	Media, ABC, NBC, CBS, television, news, network, bilingual, NoVAPJ, Spanish

online. There is a kind of asymptotic convergence — a tendency to become the same through digital form and shared network. Yet we might want to celebrate the differences in form and function for different purposes (e.g., videos are better viewed on a larger screen). Governments have the bad luck of living in interesting times of transition. Future research will help citizens and government to navigate the transition from traditional methods to emerging trends. The growing number of cities and towns that have more experience with new media wil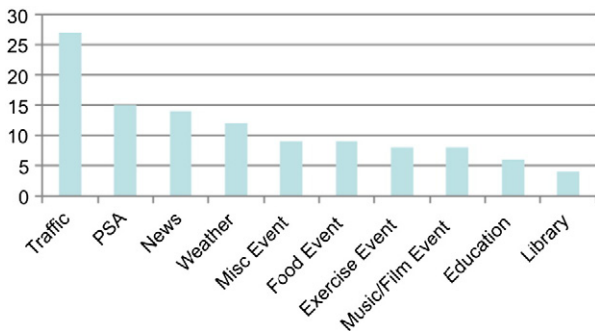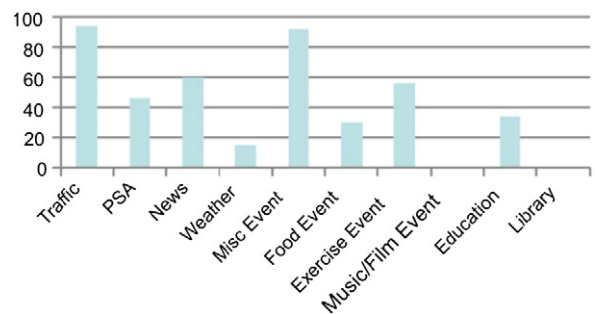l guide others in minimizing costs and pitfalls. The benefits, especially to citizens, in terms of greater access to information (e.g., searchable online video of meetings of interest) and greater sharing of concerns and ideas will lead to increased awareness, collective efficacy and civic participation. The benefits to government include pro-active problem solving and positive public relations that lead to greater political efficacy and public trust.

Acknowledgments

References

American Redcross: Social Media in Disasters and Emergencies (August 5, 2010). Retrieved Dec. 14, 2010, from http://www.redcross.org/www-files/Documents/pdf/other/SocialMediaSlideDeck.pdf.

Arlington County Blog Central. Retrieved June 4, 2010, from http://www.arlingtonva.us/departments/Communications/page70455.aspx.

Arlington County Facebook Profile. Retrieved Dec. 14, 2010, from http://www.facebook.com/ArlingtonVA.

Arlington County Flickr account. Retrieved Dec. 14, 2010, from http://www.flickr.com/photos/arlingtonva.

Arlington County News on Twitter. Retrieved Dec. 14, 2010, from http://twitter.com/arlingtoncounty.

Arlington County, VA Official Site. Retrieved Dec. 14, 2010, from http://www.arlingtonva.us/default.aspx.

Bertot, J. C., Jaeger, P. T., & Hansen, D. (2012). The impact of polices on government social media usage: Issues, challenges, and recommendations. Government Information Quarterly, 29(1), 30–40, http://dx.doi.org/10.1016/j.giq.2011.04.004.

Boyd, D., & Ellison, N. (2007). Social network sites: Definition, history and scholarship. Journal of Computer-Mediated Communication, 13(1) (article 11.).

CTRnet: Crisis, Tragedy, and Recovery network. Retrieved June 4, 2010, from http://www.ctrnet.net.

Flickr. Retrieved May 4, 2010, from http://www.flickr.com.

Furnas, G. W., Fake, C., von Ahn, L., Shachter, J., Golder, S., Fox, K., et al. (2006). Why do tagging system work? : Paper presented at the Computer-Human Interaction.

Gasbuddy: Find local gas prices. Retrieved May 4, 2010, from http://gasbuddy.com.

Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the U.S. Congress. Journal of the American Society for Information Science and Technology, 61(8), 1612–1621, http://dx.doi.org/10.1002/asi.21344.

Hughes, A. L., & Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. Gothenburg, Sweden: Paper presented at the Information Systems for Crisis Response and Management.

Hughes, A. L., Palen, L., Sutton, J., Liu, S. B., & Vieweg, S. (2008). "Site-seeing" in disaster: An examination of online social convergence. Paper presented at the International Conference on Information Systems for Crisis Response and Management (ISCRAM), Washington, DC [May 2008].

Lampe, C., Ellison, N., & Steinfield, C. (2006). A face(book) in the crowd: Social searching vs. social browsing. Paper presented at the Conference on Computer Supported Cooperative Work, Banff, Alberta Canada: Social Browsing.

O'Reilly, T. (2007). What is Web 2.0: Design patterns and business models for the next generation of software. Communications & Strategies, 65, 17–37 (1st Quarter).

Opsahl, A., 2010 County Experiments with Monitoring Social Media in Emergencies.

Protalinski, E. (2012). Facebook Has Over 845 Million Users [Retrieved from http://www.zdnet.com/blog/facebook/facebook-has-over-845-million-users/8332].

Queensland Police Service, QPS Social Media Case Study. from http://www.police.qld.gov.au/Resources/Internet/services/reportsPublications/documents/QPSSocialMediaCaseStudy.pdf.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors. WWW '10: Proceedings of the 19th international conference on World wide web (pp. 851–860). Raleigh, NC, USA: ACM.

See-Click-Fix in Arlington County. Retrieved June 2, 2010, from http://www.seeclickfix.com/arlington_2.

Smarter Transportation: 10 Social Media Tools to Navigate Your City. Retrieved May 10, 2010, from http://mashable.com/2009/09/29/social-media-transportation.

Smith, A. (2010). Government online: Pew internet & American life project.

Smith, A., Verba, S., Brady, H., & Schlozman, K. (2009). The internet and civic engagement: Pew internet & American life project.

Social Media Sells. Ogilvy & Mather Retrieved November 1, 2010, from http://www.ogilvy.com/News/Press-Releases/November-2010-Social-Media-Sells.aspx.

Tepper, M. (2003). The Rise of Social Software. netWorker, 7(3), 18–23.

Twitter blog: Measuring tweets. Retrieved Nov. 4, 2010, from http://blog.twitter.com/2010/02/measuring-tweets.html.

YouTube Blog. (2011). Retrieved from http://youtube-global.blogspot.com.au/2011/05/thanks-youtube-community-for-two-big.html.

YouTube blog (2012). Holy Nyans! 60 hours per minute and 4 billion views a day on YouTube [Retrieved Feb. 9, 2012, from http://youtube-global.blogspot.com/2012/01/holy-nyans-60-hours-per-minute-and-4.html].

Zuckerman, E. (2009). [My heart's in Accra Retrieved January 6,]. Studying Twitter and the Moldovan Protests [from http://www.ethanzuckerman.com/blog/2009/04/13/studying-twitter-and-the-moldovan-protests/].

Andrea L. Kavanaugh, A Fulbright scholar and Cunningham Fellow, Andrea Kavanaugh is Senior Research Scientist and Associate Director of the interdisciplinary research Center for Human-Computer Interaction (HCI) at Virginia Tech. Her research lies in the areas of social computing and communication behavior and effects. She has been leading research on the use and social impact of information and communication technology funded primarily by the National Science Foundation. She is the author or editor of three books; her research is also published in American Behavioral Scientist, Interacting with Computers, Journal for Computer Mediated Communication, Computer Supported Cooperative Work, and The Information Society, among others. Prior to joining the HCI Center in 2002, she served as Director of Research for the community computer network known as the Blacksburg Electronic Village (BEV) from its inception in 1993. She holds an MA from the Annenberg School for Communication, University of Pennsylvania, and a PhD in Environmental Design and Planning (with a focus on telecommunications) from Virginia Tech. She currently serves as Treasurer on the Board of the Digital Government Society (DGS), and previously served on the Board of the International Telecommunications Society (2002–08).

Dr. Edward A. Fox holds a Ph.D. and M.S. in Computer Science from Cornell University, and a B.S. from M.I.T. Since 1983 he has been at Virginia Polytechnic Institute and State University (VPI&SU or Virginia Tech), where he serves as Professor of Computer Science. He directs the Digital Library Research Laboratory and the Networked Digital Library of Theses and Dissertations. He has been (co)PI on over 110 research and development projects. In addition to his courses at Virginia Tech, Dr. Fox has taught over 74 tutorials in more than 25 countries. He has given more than 63 keynote/banquet/international invited/distinguished speaker presentations, about 154 refereed conference/workshop papers, and over 250 additional presentations.

Steven D. Sheetz is Director of the Center for Global e-Commerce and Associate Professor at the Pamplin College of Business at Virginia Tech. He received his Ph.D. in Information Systems from the University of Colorado at Boulder. His research interests include the design and use of standards for information systems development, the diffusion of software measures, computer supported collaboration systems including group cognitive mapping software and social networking for crisis situations, and information systems to support response to and recovery from crises. He has published articles in *Decision Support Systems, International Journal of Human-Computer Studies, Journal of Management Information Systems, Journal of Systems and Software, and Object-Oriented Systems.* He also holds a MBA from the University of Northern Colorado and a B.S. in Computer Science from Texas Tech University. He has substantial industry experience in database design and OO systems development.

Seungwon Yang is a PhD student in the Department of Computer Science at Virginia Tech, where he also received his B.S. degree. As a member of the Digital Library Research Laboratory under the supervision of his advisor Dr. Edward A. Fox, he explores and investigates areas of topic identification, natural language processing, ontology engineering, and machine learning applied to Web archiving and social media analysis. Other interests include a digital library curriculum development and Visual Analytics. Currently he is working on topic identification approaches and domain ontology developments as part of his doctoral study, which has connection to his funded project the Crisis, Tragedy, and Recovery Network. He has been participating in various multidisciplinary projects. He values the benefits of, and enjoys the synergies from those collaborative projects.

Lin Tzy Li is a Ph.D. student in the Institute of Computing (IC) at UNICAMP, Brazil, advised by Dr. Ricardo da S. Torres. She was a one-year PhD intern at Virginia Tech(Digital Library Research Laboratory) under the supervision of Dr. Edward A. Fox fromAug. 2010 to July 2011.She collaborated with the Crisis, Tragedy, and RecoveryNetwork project team at Virginia Tech during that period. She is also a system analystand researcher at the Telecommunication Research and Development Center (CPqD) inBrazil. Lin holds M.S. and B.S. degrees in Computer Science, and an M.B.A degree.During her Master's studies, Lin was involved in Bioinformatics of Brazilian GenomeResearch projects. Her current research interest includes GIS, Database, data fusion, andintegration of image/video, text, and geographic information retrievals. Her PhD work ison Multimodal Retrieval of Geographic Information.

Donald J. Shoemaker holds a Ph.D from the University of Georgia. He has been at Virginia Tech since 1974 and is currently Professor of Sociology. His primary research interests include youth violence and juvenile justice. He has extensive international research experience with youth deviance and juvenile justice, especially in the Philippines, but also in Turkey and Sri Lanka. In 1990, he received a Fulbright Research Scholarship to study juvenile justice in the Philippines. Recent publications include the 6th edition of Theories of Delinquency, published by Oxford University Press and Juvenile Delinquency, published by Rowman/ Littlefield. He is currently involved with the digital library research team on the Crisis, Tragedy, and Recovery (CTRnet) project, funded by the National Science Foundation.

Apostol (Paul) Natsev received the M.S. and Ph.D. degrees in computer science from Duke University, Durham, NC, in 1997 and 2001, respectively. He is currently a Staff Software Engineer and Manager in the Video Content Analysis Group at Google Research. Previously, he was a Research Staff Member at IBM Research from 2001 to 2011, and Manager of the Multimedia Research Group from 2007 to 2011. Dr. Natsev's research agenda is to advance the science and practice of systems that enable users to manage and search vast repositories of unstructured multimedia content. His research interests span the areas of image and video analysis and retrieval, computer vision, and large-scale machine learning. Dr. Natsev is an author of more than 70 publications and 18 patents (granted or pending), and his research has been recognized with several awards, including the 2004 Wall Street Journal Innovation Award, Multimedia category (for the IBM MARVEL system), 2005 IBM Outstanding Technical Accomplishment Award, 2005 ACM Multimedia Plenary Paper Award, 2006 ICME Best Poster Award, and 2008 CIVR Public's Choice Award (for the IBM IMARS system).

Lexing Xie received the B.S. degree from Tsinghua University, Beijing, China, in 2000, and the M.S. and Ph.D. degrees from Columbia University, in 2002 and 2005, respectively, all in electrical engineering. Lexing Xie is Senior Lecturer in the Research School of Computer Science at the Australian National University, she was with the IBM T.J. Watson Research Center, Hawthorne, NY from 2005 to 2010. Her recent research interests are in multimedia mining, machine learning and social media analysis. Dr. Xie has won several awards: the best conference paper award in IEEE SOLI 2011, the best student paper awards at JCDL 2007, ICIP 2004, ACM Multimedia 2005 and ACM Multimedia 2002. She also received the 2005 IBM Research Josef Raviv Memorial Postdoc fellowship in computer science and engineering.