Scalable crisis relief: Crowdsourced SMS translation and categorization with Mission 4636

Vaughn Hester
CrowdFlower
455 Valencia Street
San Francisco, CA 94103
vaughn@crowdflower.com

Aaron Shaw
UC Berkeley
Department of Sociology
410 Barrows Hall
Berkeley, CA 94720-1980
adshaw@berkeley.edu

Lukas Biewald
CrowdFlower
455 Valencia Street
San Francisco, CA 94103
lukas@crowdflower.com

ABSTRACT

Crowdsourced crisis response harnesses distributed human networks in combination with information and communication technology (ICT) to create scalable, flexible and rapid communication systems that promote well-being, survival, and recovery during the acute phase of an emergency. In this paper, we analyze a recent experience in which CrowdFlower conducted crowdsourced translation, categorization and geo-tagging for SMS-based reporting as part of Mission 4636 after a 7.0 magnitude earthquake struck Haiti on January 12, 2010. We discuss CrowdFlower's approach to this task, lessons learned from the experience, and opportunities to generalize the techniques and technologies involved for other ICT for development (ICTD) applications. We find that CrowdFlower's most significant contribution to Mission 4636 and to the broader field of crowdsourced crisis relief lies in the flexible, scalable nature of the pool of earthquake survivors, volunteers, workers, and machines that the organization engaged during the emergency response efforts.

Categories and Subject Descriptors

H.4.3 [**Information Systems Applications**]: Communications Applications; J.m [**Computer Applications**]: Miscellaneous; K.4.2 [**Computers and Society**]: Social Issues

General Terms

Human Factors, Design, Management, Languages

Keywords

Human Computation, Crowdsourcing, Crisis Response, Haiti, Translation, Geolocation, ICTD, Development

1. INTRODUCTION

Crowdsourced crisis response harnesses distributed human networks in combination with information and communication technology (ICT) to create scalable, rapid communication systems that promote well-being, survival, and recovery during the acute phase of an emergency. In this paper, we analyze a recent experience in which CrowdFlower conducted crowdsourced translation and geo-tagging for SMS-based emergency reporting as part of Mission 4636 in response to the 7.0 magnitude earthquake that struck Haiti on January 12, 2010.

Distributed, SMS-based systems for information sharing can facilitate effective responses to humanitarian crises worldwide. During the violence that followed the disputed 2007 presidential elections in Kenya, Ushahidi pioneered the use of FrontlineSMS for crisis mapping and response. Since then, Ushahidi and others have deployed FrontlineSMS and similar reporting systems in a variety of settings. One constraint of such systems emerges when linguistic and other factors prevent a scalable means of filtering, verifying, translating, and geo-tagging all the incoming SMS reports. Crowdsourcing, or the act of engaging distributed groups of people to complete microtasks or generate information, offers unique advantages in such situations when combined with (a) willing communities of volunteers and (b) a highly elastic "on-demand" labor platform.

Mission 4636 was a collaboration among technology companies, international non-governmental organizations and emergency relief agencies in the aftermath of the 2010 Haitian earthquakes. Mission 4636 combined an SMS-based reporting system with crowdsourced translation and geolocation to collect, categorize and map reports of emergency data from a distributed human network in Haiti. Many organizations and steps were involved in this process. CrowdFlower, a technology company based in San Francisco, provides a crowdsourcing plaform with rigorous quality control tools. The platform enables access to labor on demand, or to pools of distributed human judges around the world. Users of the platform create online projects, or microtasks, which are completed by multiple workers working in parallel. CrowdFlower facilitates quality-controlled completion of microtasks such as text or image moderation, sentiment analysis, and business listing verification. CrowdFlower's participation in Mission 4636 was one of the first times this crowdsourcing platform had been applied to emergency relief efforts.

This paper focuses on CrowdFlower's approach to the crowdsourcing translation and geolocation components of the broader Mission 4636 effort. In particular, we describe the system design and performance as well as lessons learned from the experience. The contributions of other organizations and the details of other parts of this group effort are, for the most part, beyond the scope of this article.

While our ability to evaluate the project was severely limited for many reasons, we find that Mission 4636's unusual capacity to engage survivors both at the stage of information collection and at the

ACM DEV'10, December 17–18, 2010, London, United Kingdom.

stage of information processing (the latter in exchange for money) has the potential to produce a triple impact by (a) alleviating emergency conditions, (b) expanding employment opportunities in the midst of a crisis, and (c) contributing to individuals' overall sense of participation in the recovery process.

Finally, we explore several ways in which the crowdsourced translation and geolocation might be improved, as well as opportunities to generalize these techniques and technologies for application in other emergency responses.

2. CROWDSOURCING AND ICTD

2.1 Crowdsourcing: Definition, Origins and Applications

We define crowdsourcing as "the act of engaging distributed groups of people to complete microtasks or generate information."[1] The new term conceals a rich history of "human computing," which David Alan Grier has traced back to early applications of calculus, including 18th century astronomy and World War I era national statistics offices whose responsibilities included the creation of ballistics tables for combat use [11]. In light of this history, we view the recent introduction of digital, networked ICTs into circuits of human computation as a catalyst for a sort of computing renaissance. With the global spread of the Internet and cellular telecommunications, the opportunities to harness the potential of distributed human intelligence and cognition have never been greater.

The increase in the number of applications of crowdsourcing follows the broader expansion of digital ICTs. The proliferation of wireless, cellular and satellite networks as well as the diffusion of devices connected through these networks facilitates the transmission of many different kinds of digital information and creates opportunities to collect, process, and analyze data on a scale previously unavailable.

In this regard, the spread of networked, digital ICTs has facilitated a rapid expansion in the tools, techniques and applications for crowdsourcing and human computation. Several online labor markets and commercial crowdsourcing service providers, including (among the U.S. based firms) LiveOps, Amazon.com's Mechanical Turk, Casting Words, and CrowdFlower (a) contribute to the growth of the tools, techniques, and applications for crowdsourcing and (b) constitute an emergent sector within the ICT industry.

The analysis and application of crowdsourcing has grown in a similar manner. Among many noteworthy examples in the computational and social sciences, Von Ahn and collaborators have developed widely adopted image labeling, text transcription, and music tagging tools that harness distributed human intelligence [27, 19, 26]. Also, a range of research applications involving the use of crowdsourcing to conduct behavioral and labor experiments [16, 15, 6, 5], filter inexpert judgments [24], conduct user studies [18] and analyze large bodies of textual content [13] have introduced crowdsourcing into diverse disciplinary contexts.

The rise of crowdsourcing applications has elicited divided opinions among researchers about the ethical, economic and political implications of these phenomena. Some scholars have emphasized the potential for crowdsourcing and related practices to bring about social welfare gains by facilitating (a) innovation, (b) distributed problem solving and (c) access to information resources[2]. Others have made the case that crowdsourcing (as an example of human computing) also makes it easier to exploit workers, conduct malicious tasks, or generally devalue human intelligence[28].

We contend that crowdsourcing can and will produce positive social welfare benefits in a variety of contexts, but that nothing in the technologies or techniques makes this inherently true. We support this position through our analysis of Mission 4636, which entailed the confluence of several different actors, technological platforms, networks, and circumstances following the 2010 earthquakes in Haiti.

2.2 Relevance to ICTD

Mission 4636 exemplifies the growing number of projects that incorporate crowdsourcing and SMS networks to advance the social welfare aims at the core of ICT for development (ICTD) research.[2] In this sense, Mission 4636 extends many existing tools and approaches while also contributing several unique innovations into the fields of ICTD and emergency response. We review a few salient examples of ICTD and emergency response projects that use crowdsourcing before turning to our analysis of Mission 4636 below.

Efforts to incorporate crowdsourcing technologies and practices into ICTD projects build on existing work with networks of various kinds. For example, the use of sensor networks to monitor water quality or irrigation levels also harnesses distributed, large-scale information gathering to comparable effects, but it involves a less continuous flow of human participation than most crowdsourcing work[22, 23].

Crowdsourcing has begun to play a more prominent role in development initiatives and emergency response efforts. An ideal example of this trend, the San Francisco nonprofit Samasource facilitates digital work opportunities for underserved groups such as women, children, and refugees around the world. As part of a well-publicized 2007 search and rescue effort, emergency workers collaborated with users of Google Earth and Amazon Mechanical Turk to look for Steve Fossett, a pilot whose plane had disappeared over the Nevada desert.[3] While the crowdsourced search ultimately failed, we find no reason to believe that the failure will prevent similar attempts in the future and find the effort to achieve a large-scale search and rescue operation using a combination of paid and unpaid volunteers especially relevant to the case of Mission 4636.

Similarly, SMS-based crowdsourcing projects represent a relatively new addition to the existing repertoire of ICTD applications for mobile phones. SMS offers numerous practical advantages in regions with high cell phone diffusion rates and GSM network coverage such as many parts of the Caribbean[7, 14]. The txteagle project represents the most innovative effort to use SMS-based crowdsourcing to enable cell phones subscribers in Kenya and Rwanda to earn income in exchange for completing microtasks[8].

Rapid emergency response adds an additional layer of complexity and difficulty to ICTD projects and research[20]. Ken Banks and the creators of FrontlineSMS first applied an SMS-based reporting system to crowdsource election monitoring in Nigeria in 2007[1]. Subsequently, the FrontlineSMS system was extended

[1]This definition builds on the original usage by Jeff Howe, who coined the term[17].

[2]Brewer et al. provide an early, synthetic overview of ICTD[3]. Following recent work by Burrell and Toyama, we refrain from attempting to narrowly define ICTD or development, but use the terms here to refer to the scholarly community that (a) identifies itself with similar terms and (b) focuses its research on the social impact and implications of ICTs in geographic regions historically considered to be "developing" or "underdeveloped"[4].

[3]More information about the search and links to news stories can be found at: http://www.stevefossett.com/

by Ushahidi to help map post-election violence in Kenya later that year. FrontlineSMS itself incorporated this system into the FrontlineSMS:Medic project to promote health care through effective data collection and patient-physician interactions.[4]

Building on previous ICTD, crowdsourcing, emergency response, and SMS-reporting projects, the key innovations of Mission 4636 consist of several distinct components. First, the project expands on the efforts of FrontlineSMS, Ushahidi and related groups to integrate both the survivors of emergencies as well as the promoters of ongoing community development into the recovery process. Second, Mission 4636 also takes advantage of the distributed effort of volunteers and paid workers to achieve a response of sufficient scale and speed in the wake of an unexpected event. Finally, and perhaps most uniquely, Mission 4636 represents an attempt to bypass the dual barriers of scale and language through a complex, multi-step workflow that incorporated a number of organizational actors around the world. Such an adoption of multiple tools as well as collaborative strategies in response to a sudden crisis are becoming a viable and valuable resource for emergency responders[10].

3. CROWDSOURCING CRISIS RESPONSE IN HAITI

3.1 January 2010 Earthquakes

On January 12, 2010, a magnitude 7.0 earthquake struck Haiti 25 miles west of Port-au-Prince. In the 12 days following this quake, 52 aftershocks of magnitude 4.5 or higher were recorded.

Despite massive devastation of basic infrastructure in Port-au-Prince as well as complete or periodic disruption of many essential services, most of Haiti's cell tower infrastructure remained intact.

Over 20 countries sent military personnel to contribute to relief efforts on the ground. Dozens of international NGOs and agencies deployed staff and resources to Haiti as well. Communication, coordination and oversight were chaotic in the first few days after January 12.

3.2 Mission 4636

Within 48 hours of the first earthquake, Josh Nesbit of FrontlineSMS:Medic and Katie Stanton of the U.S. State Department convinced DigiCel, the largest telco in Haiti, to set up a short code – "4636" – that any individual could text for free to request help. Robert Munro of Energy for Opportunity and Brian Herbert of Ushahidi set up a workflow where Kreyol-speaking volunteers could translate the messages into English and classify them. The translated messages were then routed to aid agencies and workers who could send relief. Once the system was working, InSTEDD (in collaboration with Thompson Reuters) worked on the ground to broadcast the existence of the "4636" short code to as many Haitians as possible using radio and other accessible forms of communication.

3.3 Distributed Tagging and Translation with Crowdflower

People throughout Haiti used this short code system to send very urgent messages and pleas for help. As the volume of messages grew, so did the need for a more robust workflow platform. CrowdFlower began receiving feeds of SMS messages via Ushahidi (and its Haitian telco partners) and facilitating their review. Volunteers on the CrowdFlower platform provided translation, categorization and geo-tagging.

[4]A case study by Goldstein and Rotich situates Ushahidi's early work within the broader context of the Kenyan elections[9].

Each worker was shown an SMS message from Haiti. The worker was asked whether he or she could understand and translate the message. If not, the worker had the option of moving onto another message. A second question allowed the worker to flag a particular message if it was indecipherable or if it did not contain any information.

The worker then completed as many fields as possible in response to questions about the SMS message.[5]

- Sender's first and last name (free text entry)
- Sender's gender (radio button)
- Sender's status: missing, survived, wounded, deceased or other
- Category of the aid request. The seven main categories were Emergency, Threats, Vital Logistics, Response, Other, Persons News and Child Alone.
- Zone or neighborhood of the request/report

Once the messages had been processed and categorized using this framework, CrowdFlower posted feeds of translated and tagged messages.

4. RESULTS

4.1 Worker Population

Kreyol speakers around the world were asked to contribute to the translation and tagging efforts. Through word of mouth and a small amount of viral online marketing about Mission 4636, the pool of volunteer translators grew throughout the Haitian diaspora. Mission 4636 also created a short viral video to publicize its efforts.[6]

The Kreyol SMS translation and classification work was completed with contributions from a large number of volunteers around the world. In Haiti, residents also played an integral role in the SMS-based relief efforts. Before the first earthquake, Samasource launched a new work center in Haiti, which now made it possible to hire earthquake survivors to contribute to crowdsourced relief efforts. This Samasource service partner assumed a large amount of the earthquake relief responsibilities, not only providing labor for the emergency message routing, but also creating jobs for the local economy.

We report geocoded IP information on the locations of the workers who translated and tagged SMS messages via CrowdFlower in Table 1.

The distribution of workers' IP addresses provides an approximate sense of the geographic distribution of the workers themselves. The vast proportion of U.S. addresses illustrates that these crowdsourcing efforts were dominated by volunteers and workers living outside the earthquake affected region.

4.2 System Throughput

The combined effort of the workers, volunteers and various partner organizations involved produced low latency and accessible results. CrowdFlower processed approximately 100,000 SMS messages in total and collected over 16,000 judgments (translations) on the messages that were decipherable and legible. At peak volume, CrowdFlower workers and volunteers filtered over 5,000 SMS

[5]A preview of the task is available through the Crowdflower internal judgment interface: http://www.crowdflower.com/judgments/mob/4980.

[6]Available at http://www.youtube.com/watch?v=QVvYauvydac

Table 1: Worker Location by Country (geocoded IP address)

	Count	Percentage†
USA	11616	89.49
Canada	853	6.57
Haiti	232	1.79
Switzerland	155	1.19
Colombia	52	0.40
Germany	39	0.30
France	16	0.12
Guadeloupe	10	0.08
Japan	4	0.03
Great Britain	2	0.02
Kenya	1	0.01
Total	**12980**	**100.00**

† No IP information was available for 9289 workers or 41.7% of the aggregate workforce.

messages in one hour. The average response time to filter, translate, map/geocode and categorize a message did not exceed two minutes. Parts of the emergency SMS message feeds – and the maps generated by Ushahidi – are now used by a growing number of organizations, including the Red Cross, Plan International, charity:water, the U.S. State Department, International Medical Corps, AIDG, USAID, FEMA, the U.S. Coast Guard Task Force, World Food Program, SOUTHCOM, OFDA and UNDP.

Figure 1 illustrates the overall throughput of the system as the number of individual workers participating in Mission 4636 through the Crowdflower platform per day during the months immediately following the earthquakes. The graph begins January 28, 2010 and ends May 12, 2010.

Figure 1 also conveys the uneven spikes typical of many crowdsourcing jobs. We are not sure what caused the unlikely spike around mid April. To reiterate, the ability to accomodate unpredictable contractions and expansions in work volume and the related volume of workers needed to complete this work is a tremendous advantage of dedicated, crowdsourcing systems.

A key component of the task consisted in categorizing the messages that came through. Table 2 provides summary information about the intelligible aid request messages that were categorized by workers and volunteers through the CrowdFlower platform. The categories are sorted by prevalence.

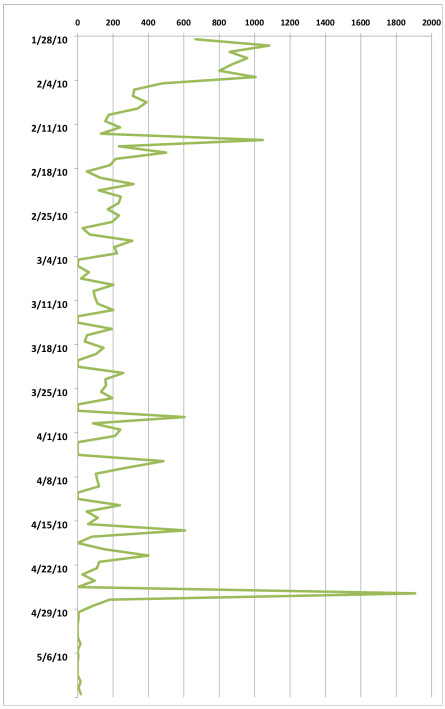
Note the prevalence of the "Other" and "Blank" categories, which made up over 80% of the total. Also note that despite the high number of blanks, "the crowd" applied a total of 12,364 "non-blank" categories, or 55.52% of the total.

4.3 Quality Control

The urgent nature of Mission 4636 did not allow for the application of redundancy that is a traditional step in CrowdFlower quality control. In other crowdsourcing projects, multiple workers respond to the same question, and an aggregation method (such as majority voting) then produces a robust estimate of the single accurate response. In the Mission 4636 efforts, the majority of the SMS messages were reviewed by a single judge only. Accordingly, it was not possible to calculate precision or accuracy by measuring agreement across multiple workers on a single SMS message.

Despite the absence of multiple judgments on each incoming

Figure 1: Traffic per day (workers)



SMS, CrowdFlower employed several strategies for quality control. A tactic for maximizing the skills of Kreyol speakers was to filter out illegible, indecipherable or nonsense messages. By allowing workers to flag and remove these SMS data units, overall latency was streamlined and translators remained focused on substantive messages to translate and categorize.

Mission 4636 also established a chat room for the online workers to facilitate collaboration and clarification. This parallel stream of communication allowed workers to confirm and share additional details on exact locations and addresses in Haiti. We believe this communication contributed to heightened accuracy and more efficient throughput.[7]

Once an SMS message was processed, translated and categorized by CrowdFlower volunteers, it was immediately pushed to an outbound feed to Ushahidi. During the project, a CrowdFlower effort was made to improve the quality of the processed SMS messages by improving the speed and accessibility of the outbound feed of the translated and categorized messages. CrowdFlower and Mission 4636 published the feed with the aim of making the messages available to as many relief agencies and actors as possible.

5. DISCUSSION

5.1 Impact

One challenge of operating in an acutely affected environment after a major emergency is the difficulty of accurately assessing the impact of many emergency response efforts. This was complicated by the fact that the CrowdFlower SMS translation and categorization effort employed workers from all over the world and served primarily to distribute information across dozens of separate organizations and agencies operating in Haiti. We were unable to collect

[7] A brief report by Robert Munro posted on the Ushahidi blog supports this claim[21].

Table 2: Summary of SMS Categorization Results

Category	Count	Percentage
Blank†	9905	44.48
Other	9204	41.33
Response request	1435	6.44
Vital logistics	734	3.30
Emergency	691	3.10
Persons news	235	1.06
Threats	52	0.23
Child alone	13	0.06
Total	**22269**	**100**

† These messages were analyzed, but category information was left blank.

systematic data to evaluate CrowdFlower's role in Mission 4636 in a rigorous way. In the absence of such data, we discuss the impact of the project in this section with the caveat that our findings are not conclusive and may diverge from the observations of other individuals and organizations involved in the relief effort.

As described in the previous section, overall system performance scaled well to match the flow of incoming messages. The massive number of volunteers worldwide in combination with the Haiti-based workforce brought online by Samasource meant that the system processed most emergency messages with extremely low latency. For messages like "Non mwen se lucaint lucoit madanm mwen ansent li rive le poul akouche nou nan delma 31 ri maryen n 21 nan lakou legliz apostolik anfas site jeremi, mpa" ("condition bloody about. undergoing children delivery corner of delmas 31 and rue marine") it was crucial not just to be fast, but to have local knowledge to get the exact longitude and latitude from an ambiguous 140 character message as well as an accurate classification so that the right aid agency can be deployed. In the case of this particular SMS, USGS responded "just got emergency SMS, child delivery, USCG are acting, and, the GPS coordinates of the location we got from someone of your team were 100% accurate"[12].

The organizations receiving and using Mission 4636 information to direct relief services provided strong, positive feedback. Craig Clark of the U.S. Marine Corps commented on the text message project: "I wish I had time to document to you every example, but there are too many and our operation is moving too fast ... I say with confidence that there are hundreds of these kinds of [success] stories. The Marine Corps is using your project every second of the day to get aid and assistance to the people that need it most." Many other relief actors, including the State Department, shared similar comments on the utility of Mission 4636 and the unprecedented efficiency it added to relief efforts[12].

Another consideration with respect to the overall project impact was the speed with which it was deployed. The SMS translation project was developed and launched extremely quickly, and there were numerous challenges throughout this process. Publishing an RSS feed of the translated SMS that was accessible to many different groups was key in allowing aid agencies to start downloading the particular categories of messages that were important to their work. Ushahidi's maps of the reported incidents and the requests for help made an important delineation of the overall trends separate from the individual needs.

5.2 Limitations and Recommendations

After evaluating the lessons learned from CrowdFlower's experience with Mission 4636, we found a few key limitations and recommendations.

The biggest technical constraint had to do with the various SMS providers changing the formats of their feeds during the aftermath. Such changes imposed major obstacles for SMS message processing, as CrowdFlower engineers had to alter the underlying structure of the message-processing system in a hurry after each change. We recommend that telecommunications providers do everything in their power to maintain consistent data formats during emergency situations.

As described above, CrowdFlower was constrained in its ability to implement a more robust set of quality control tools in this case. Previous research has demonstrated the significant gains in judgment accuracy and precision that can result from even relatively simple quality measurement techniques[25, 24]. CrowdFlower deploys many of these same techniques with its clients through its online platform; however, the Mission 4636 tasks did not incorporate many of these tools as a result of the urgency and speed with which the response system was deployed.

This experience reflects the fact that in crowdsourcing, as in many computational endeavors, we often encounter the tension between precision and recall. In the case of Mission 4636, the crowdsourcing system was optimized for recall, perhaps at the expense of precision.

In future iterations of crowdsourced microtasks to support disaster relief, it would also be advantageous to add flexible functionality to the CrowdFlower platform. Specifically, certain categories or types of SMS messages may warrant review by more than one human judge. This dynamic or variable redundancy on highly urgent messages could reveal more details about the actual location of the message and prevent any mistakes that a single translator may have made.

Another lesson learned concerns the design and presentation of the SMS data itself through the CrowdFlower worker interface. The task display included information about the number of SMS messages awaiting judgment. If this number was too big, volunteers reported that they felt hopeless, yet others reported that the transparency surrounding the pending workload provided a great motivator for the people churning through the tasks. We draw two conclusions from these reports. First, we might make the display of the queue an optional feature, allowing workers and volunteers the latitude to hide it if they wish. Another possible improvement would be to pre-populate the task with a Google translation of the SMS text to ease the workload of the translators, thereby enhancing the speed of the system and reducing frustration simultaneously. Workers could simply correct or verify the machine translations, pointing out any specific or missing details that need enhancement.

A final concern regards the overall quality of the information flowing through the system. Volunteered information and crowdsourced analysis may produce results of questionable precision and accuracy. At the same time, the effective integration of such "imperfect signals" can complement existing emergency response systems. In other words, emergency information that may be imperfect is better than an absence of any information in the context of the acute phase of an emergency cycle. An evaluation of this claim should be left to future work. As in all crowdsourcing endeavors, the incentives of the participants will affect the end result of the project. Disaster victims in dire situations are, in our opinion, less likely to attempt to scam a lifesaving system than unethical online workers in an enterprise crowdsourcing task may be.

6. FUTURE WORK

In future crowdsourced emergency response efforts, we propose adding layers of quality control along the lines laid out in previous research[25, 24]. Given the success of this project in Kreyol-English translation, we also can engage Kreyol speakers to verify the quality of translations. This creates a multi-step workflow to validate results, which would increase completion time and reduce throughput. Our hypothesis is that this would generate substantial improvements in quality.

In future applications of crowdsourcing in disaster relief, more nuanced results could be available through more use of the CrowdFlower platform. Specifically, having more people review or verify certain SMS messages (i.e., the more urgent ones) could allow workers to extract more details about the actual location of the message. Multiple judgments per message would prevent any mistakes that a particular single translator may have made. [8]

It is clear that new technology and cross-platform collaboration stand to make significant contributions to future disaster preparedness and relief efforts. Crowdsourcing has a role within these efforts and should be considered a highly flexible, highly responsive tool for any and all actors. As mobile phone penetration grows around the world, we can anticipate growth in the coverage of SMS-based relief efforts in the event of an emergency. As Internet penetration grows, we anticipate an increase in the available worker pools to assist with crowdsourced translation, categorization and tagging efforts to support on-the-ground relief efforts. Finally, as governments, nonprofits and humanitarian relief agencies learn more about the possibilities that lie in this new tool for emergency relief, we are confident that greater numbers of lives will not only be changed but may very well be saved.

Ongoing efforts by CrowdFlower and others are focused on establishing a more permanent version of this short code system and SMS-based reporting for hurricane season in Haiti. International relief agencies and researchers should examine the feasibility of integrating crowdsourcing into crisis response assessment methods and toolkits. Crowdsourcing can complement existing emergency response systems and networks, but greater information and education on crowdsourced options and approaches will need to be circulated among the international relief community. Local authorities around the world should also look to their version of 911 systems to see whether parts of the response cycle could be crowdsourced efficiently. Situations that are less urgent than natural disasters can create spikes in the volume of digital information that enters these automated systems. Crowdsourcing can facilitate processing such unexpected spikes.

A major danger of expanding crowdsourced emergency response stems from the potential for poor integration with existing systems. In the aforementioned example of the search and rescue efforts to find Steve Fossett, Mechanical Turk workers produced a vast number of false-positives, straining emergency agencies' resources and willingness to collaborate.

Future projects should evaluate the effectiveness of these kinds of solutions as well as a more comprehensive evaluation of their integration and utility in acute emergency repsonse. Future research should also assess SMS-reporting system performance as well as short and long-term impacts on survivor populations and future emergency preparedness and response efforts.

[8]As of the time of publication of this paper, CrowdFlower was involved with the Pakreport project, a similar SMS-based relief collaboration with Ushahidi and other agencies. In this instance, there was an opportunity to draw upon the lessons learned with Mission 4636 and to structure the SMS processing such that each message was moderated by three volunteers. More details on the broader Pakreport effort can be found at http://www.pakreport.org/ushahidi/

7. CONCLUSIONS

We have presented an overview of CrowdFlower's efforts to maximize the quality of the crowdsourced information geolocation and categorization as part of Mission 4636. Our focus in this analysis on evaluating tools and techniques rather than project participants' interpretations therefore positions these findings closer to the "problem solving" pole of the pragmatist-interpretivist spectrum of ICTD research laid out in a recent essay by Burrell and Toyama[4].

We find that CrowdFlower's most significant contribution to Mission 4636 and to the broader field of crowdsourced crisis relief stems from the scalability of the pool of earthquake survivors, volunteers, workers, and machines that the organization engaged at an acute stage of the emergency response effort. This collaboration generated a high volume of filtered, categorized, and geocoded messages that were used by relief workers and agencies to direct assistance where it was needed most.

While we are unable to conduct a more thorough review of the project impact in Haiti among earthquake survivors and their families, our analysis offers several significant take-away points:

- The importance of task design.
- The importance of effective collaboration with other organizations and actors involved.
- The opportunities made possible by efficient, large -scale information processing.
- The opportunities to improve system throughput and quality control.

Mission 4636 came to life because multiple organizations were able to work together quickly and effectively. It illustrated the advantages of a flexible, crowdsourced workflow for disaster relief management. Businesses use crowdsourcing because it eliminates the need to forecast or anticipate work capacity in advance. Disaster relief is simply an extreme extension of this same situation. Traditionally, harnessing thousands of volunteers would represent a logistical quagmire, but crowdsourcing platforms are specifically suited for these types of amorphous, virtual labor forces. Through viral channels and a microtask framework, it was possible (and will be possible in the future) to have thousands of people around the world doing mission-critical work within hours or days.

8. ACKNOWLEDGMENTS

The authors are grateful for the support and collaboration of everyone at CrowdFlower and the participating organizations in Mission 4636. We especially thank Anisha Sekar and Diyang Tang (CrowdFlower interns) for their invaluable assistance with the preparation of this manuscript.

9. REFERENCES

[1] K. Banks. And then came the nigerian elections: The story of frontlinesms. *SAUTI: The Stanford Journal of African Studies*, (Spring/Fall):1–4, 2007.

[2] Y. Benkler. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press, 2006.

[3] E. Brewer, M. Demmer, B. Du, M. Ho, M. Kam, S. Nedevschi, J. Pal, R. Patra, S. Surana, and K. Fall. The case for technology in developing regions. *IEEE Computer*, 38(6):25–38, 2005.

[4] J. Burrell and K. Toyama. What constitutes good ictd research? *Information Technologies International Development*, 5(3), Fall 2009.

[5] D. Chandler and A. Kapelner. Breaking monotony with meaning: Motivation in crowdsourcing markets. *University of Chicago mimeo*, 2010.

[6] D. L. Chen and J. J. Horton. The wages of paycuts: Evidence from a field experiment. *Working Paper*, 2009.

[7] H. Dunn. Teleworking the mobile caribbean: Emerging patterns of broadband-assisted remote work among the marginalized in jamaica and trinidad and tobago. *Information Technologies International Development*, 5(2), 2009.

[8] N. Eagle. txteagle: Mobile crowdsourcing. In *Internationalization, Design and Global Development*, pages 447–456. Springer, Berlin, 2009.

[9] J. Goldstein and J. Rotich. Digitally networked technology in kenya's 2007-2008 post-election crisis. *Berkman Center for Internet and Society: Internet & Democracy Case Study Series*, 2008.

[10] M. F. Goodchild and J. A. Glennon. Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, 2010.

[11] D. A. Grier. *When Computers Were Human*. Princeton University Press, 2005.

[12] V. Hester. Crowdflower impacts the future of disaster relief. http://haitirewired.wired.com/profiles/blogs/crowdflower-impacts-the-future.

[13] D. Hopkins and G. King. A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, 54(1):229–247, 2010.

[14] H. Horst and D. Miller. *The Cell Phone: An Anthropology of Communication*. Berg Publishers, 2006.

[15] J. Horton and L. Chilton. The labor economics of paid crowdsourcing. *Proceedings of the ACM Conference on Electronic Commerce 2010 (forthcoming)*, 2010.

[16] J. J. Horton, D. Rand, and R. J. Zeckhauser. The online laboratory: Conducting experiments in a real labor market. *SSRN eLibrary*, 2010.

[17] J. Howe. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. Crown Business, 2008.

[18] A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. *Proceedings of Computer Human Interaction (CHI-2008)*, 2008.

[19] E. Law, L. Von Ahn, R. Dannenberg, and M. Crawford. Tagatune: A game for music and sound annotation. In *International Conference on Music Information Retrieval*, pages 361–364. Citeseer, 2003.

[20] C. F. Maitland, N. Pogrebnyakov, and A. F. van Gorp. A fragile link: Disaster relief, icts and development. In *Information and Communication Technologies and Development, 2006*, 2006.

[21] R. Munro. The nuts and bolts behind 4636 in haiti. http://blog.ushahidi.com/index.php/2010/01/22/the-nuts-and-bolts-behind-4636-in-haiti/.

[22] J. Panchard, S. Rao, T. Prabhakar, H. Jamadagni, and J. Hubaux. Common-sense net: Improved water management for resource-poor farmers via sensor networks. In *Information and Communication Technologies and Development, 2006.*, 2006.

[23] N. Ramanathan, L. Balzano, D. Estrin, M. Hansen, T. Harmon, J. Jay, W. Kaiser, and G. Sukhatme. Designing wireless sensor networks as a shared resource for sustainable development. In *Information and Communication Technologies and Development, 2006.*, 2006.

[24] V. S. Sheng, F. Provost, and P. G. Ipeirotis. Get another label? improving data quality and data mining using multiple, noisy labelers. *Knowledge Discovery and Data Minding 2008 (KDD-2008)*, 2008.

[25] R. Snow, B. O'Connor, D. Jurafsky, and A. Y. Ng. Cheap and fast—but is it good? evaluating non-expert annotations for natural language tasks. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2008)*, 2008.

[26] L. Von Ahn. Games with a purpose. *IEEE Computer Magazine*, 39(6):96–98, 2006.

[27] L. Von Ahn and L. Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 319–326. ACM, 2004.

[28] J. Zittrain. Ubiquitous human computing. *Philosophical Transactions of the Royal Society A*, 366:3813–3821, 2008.