

INTERNEWS CENTER FOR INNOVATION & LEARNING

innovation.internews.org

MAPPING THE MAPS

A Meta-Level Analysis of Ushahidi & Crowdmap

CROWDGLLOBE

A Report form the Internews Center for Innovation and Learning

July 2012, Washington DC

CrowdGlobe.net

THE AUTHORS AND RESEARCH TEAM

Catie Bailard is Assistant Professor of Media and Public Affairs in the School of Media and Public Affairs at George Washington University. Before joining the SMPA faculty in 2009, Catie received her doctorate in political science from UCLA. She graduated with a 3.947 cumulative GPA with concentrations in American Politics, Formal and Quantitative Methods, and International Relations. Throughout Catie's academic career, her research agenda has primarily focused on the intersection of politics and information and communication technologies. This fascination with the effect of media on political behaviors and outcomes began in college as a major in UCLA's Communication Studies Department, a top-ranked undergraduate department, where she graduated *cum laude*. It was this experience that inspired Catie's decision to pursue a doctoral degree in political science at UCLA, which provided her with a broad substantive understanding of political science and political communication, as well as rigorous training in methodology. Catie's primary research focus is the cross-national analysis of the Internet's influence on people's evaluations and expectations of their governments, particularly focused on individuals' satisfaction with how democracy functions in their own nations. Other recent work includes an analysis of the impact of mobile phones on corruption in Africa (published in *Political Communication*).

Rob Baker has over ten years of experience as a web and new media developer, trainer, and manager as Project G Outreach Manager, responsible for documentation, toolkits, and working with clients. Before officially joining Ushahidi, his contributions to their community earned him the first ever induction to the Trusted Developer Network for his work as technical or project lead on dozens of Ushahidi deployments from crisis response to civic engagement around the world as well as the creation of the Ushahidi Community website. In addition, he acted as Director of the Universities for Ushahidi program, a 2011 initiative to train students from around the world on mobile and mapping technology. Baker is also a member of the Humanitarian OpenStreetMap Team, writes and develops mobile-ready textbooks, and teaches online courses on ICT for humanitarians through TechChange. Rob has coordinated international educational programs and media projects in Africa, Haiti, and the Middle East, as well as technical development of online projects within non-profit organizations. He has contributed to the technical development of several open-source projects and codebases, while he has also produced, edited, and shot short form video for the web with several feature humanitarian aid pieces. Baker is a US Delegate of the U.S.-Russian Bilateral Presidential Commission Subgroup on Mass Media. He is based in Washington, DC and can also be found at @rrbaker.

Matt Hindman is Associate Professor in the School of Media and Public Affairs at George Washington University. His work focuses on political communication with a concentration on Internet politics. Dr. Hindman's book *The Myth of Digital Democracy*, published in 2009 by Princeton University Press, looks at the Internet's impact on American politics. The book won Harvard's Goldsmith Book Prize as well as the Donald McGannon Award for communication research. In the past two years, Dr. Hindman has been given presentations or invited lectures at Harvard, Yale, Princeton, Columbia, the University of Pennsylvania, Stanford, and Oxford. The book has been referenced by members of the Federal Communications Commission in public speeches, and featured on NPR's *On the Media*. In addition to the book, Dr. Hindman has published on other topics including online campaigning, "open source" politics, and the online public sphere. His article "The Real Lessons of Howard Dean" was deemed the best article of 2006 by the Information Technology and Politics section of the American Political Science Association. Hindman has published several op-eds in the *New York Times* on technology issues. Dr. Hindman

has previously been an assistant professor of political science at Arizona State University. For the 2010-2011 academic year, he will be a non-resident Faculty Associate with the Berkman Center for Internet and Society at Harvard. His next book will be on the political economy of the online public sphere.

Steven Livingston is Professor of Media and Public Affairs and Media and International Affairs with joint appointments in the School of Media and Public Affairs and the Elliott School of International Affairs. He holds a PhD in political science from the University of Washington (1990). In addition to his teaching and scholarship, he has held a variety of administrative posts at GW, including stints as the director of the School of Media and Public Affairs and director of the Political Communication Program. He is also the founding director of the Institute for Public Diplomacy and Global Community (formerly the Institute for Public Diplomacy). His current research considers new forms of technologically enabled collective action in pursuit of public goods where the state is incapable or unwilling to meet these needs. His most recent book, *When the Press Fails* (University of Chicago Press), looks at media and public opinion dynamics prior to and in the early stages of the Iraq war. His earlier work on global real-time media coverage of war and conflict is the foundational scholarly work in this area. His work on collective action, governance, and information and communication technology has taken him to some thirty countries in the last four years, including multiple trips to Iraq and Afghanistan, as well as stays in several African countries, India, Malaysia, and to South and Central America.

Patrick Meier (PhD) is an internationally recognized thought leader on the application of new technologies for crisis early warning and humanitarian response. He presently serves as Director of Social Innovation at the Qatar Foundation's Computing Research Institute (QCRI) where he spearheads applied research in advanced computing to develop next-generation humanitarian technologies. Prior to QCRI, Patrick co-founded and co-directed the Harvard Humanitarian Initiative's (HHI) Program on Crisis Mapping & Early Warning and served as Director of Crisis Mapping at Ushahidi. He also co-founded the CrisisMappers Network, Standby Volunteer Task Force and the Digital Humanitarian Network. Patrick holds a PhD from The Fletcher School, a Pre-Doctoral Fellowship from Stanford and an MA from Columbia University. He blogs at iRevolution and tweets at @patrickmeier.

ACKNOWLEDGMENTS

This report was a collaborate effort, drawing on the expertise of a wide range of contributions from Internews, George Washington University (GWU) and Ushahidi. Gregory Asmolov, a doctoral student at Media and Communications Department, London School of Economics, first conceptualized CrowdGlobe as an online research hub for crowdsourcing deployments and we are grateful to him for inspiring its creation and his continued contributions. The Internews Center for Innovation & Learning catalyzed the project; GWU faculty carried out the analysis of the data provided by Ushahidi. Nikki Usher (GWU), Amanda Noonan (Internews) and Mark Frohardt (Internews) provided important guidance and feedback throughout the drafting process of this report. Many thanks to Brian Herbert (Ushahidi) for sharing the Crowdmap data without which this report would not exist. Finally but certainly not least, GWU Research Assistants coded a considerable amount of data that made the empirical analysis for this report possible.

CREDITS

Graphics: iHub Research

Design: Kirsten Ankers, CitrineSky Design

CONTENTS

Executive Summary	3
Introduction.....	5
Ushahidi & Crowdmap	6
Research Questions And Design.....	8
Quantitative Results & Analysis	10
Power Law Distribution	11
Survey Research	14
Survey Results & Analysis	15
Challenges.....	17
Successes.....	17
Conclusions And Recommendations.....	19
Case Studies	21
Case Study 1: Haiti	21
Case Study 2: Libya.....	23
Case Study 3: Japan.....	24
Case Study 4: Sudan.....	25
Case Study 5: Egypt.....	26
Appendix I – Dictionary For Crowdmap Data.....	28
Appendix Ii – Survey Research Questions	30

EXECUTIVE SUMMARY

The purpose of the CrowdGlobe project is to study various crowdsourced-mapping platforms, searching for data patterns that can tell us more about the functions of these tools and their limits as well as potentials. The CrowdGlobe.net website is an integral part of the CrowdGlobe project, providing researchers with additional case studies, meta-level datasets and analysis. CrowdGlobe is strictly platform agnostic and seeks to analyze all crowdsourced mapping technologies.

For this, its first case study, CrowdGlobe has analyzed Ushahidi and Crowdmap data as well as these platforms' user base. The Ushahidi platform, which means, "witness" in Swahili, is a free and open source tool that integrates information collection features with a live map. Ushahidi, the company, subsequently launched Crowdmap, a hosted version of the Ushahidi platform, which is easier to use since downloading the software and installing it is not necessary. When the CrowdGlobe research project was launched in October 2011, a total of 12,795 Crowdmaps had been created in over 100 countries. This presented CrowdGlobe researchers with an ideal first use-case for the project. The aim of this first report is to develop a better understanding of how Crowdmap (and Ushahidi) have been used and to analyze the data they have generated over recent years.

Our work took advantage of statistical analysis, quantitative content analysis and exploratory surveys. The quantitative analysis revealed that 93% of the 12,000+ Crowdmaps analyzed had fewer than 10 reports while 61% of Crowdmaps were identical to the default Crowdmap setting, i.e., they had not been customized or used at all. This "long tail" distribution of Crowdmaps follows a power law distribution, a common feature in many online platforms, as well as in a number of occurring phenomena. Crowdmaps with 21 to 10,000 reports were selected for further analysis, resulting in a data set of 585 maps. About 30% of these focused on North America while 18% focused on Western Europe and 16% on Africa. On average, these Crowdmaps had 814 reports but the median number of reports for this set of deployments was substantially lower, which is not surprising considering that Crowdmaps follow a power law distribution.

When the analysis is broken down by region, the relative frequency with which themes emerged in the regional deployments differed dramatically. For example, taking into account recent events, it is not surprising that the most common themes that emerged from the 63 deployments in the Middle East and Northern Africa pertain to: crime and public safety issues (43%), human rights abuses (40%), emergency-related infrastructural issues (30%), and political organization (25%). The distribution of themes in the 79 Western European deployments, on the other hand, paints a very different picture, with entertainment and leisure appearing in 32% of the deployments, followed by non-emergency infrastructural issues (25%), and media reports (23%). It is also not surprising that the 16 deployments from the Caribbean region, 12 of which hailed from Haiti specifically, heavily featured issues related to the occurrence and aftermath of a natural disaster (63% and 50%, respectively), emergency-related infrastructural issues (63%), health and medical-related issues (50%), and crime and public safety issues (38%).

In addition, surveys were sent to all 12,795 Crowdmap users to better understand how they used the platform and to assess their experience. About 80% of respondents are men and the average age of a user is 40 years old. As for educational background, an impressive 43% of users have a post-graduate degree and a total of 84% have at least a college degree. Approximately 53% of users responded as having no prior experience in using online mapping technologies, which is not surprising since the technology is still relatively new. The most important reason why Crowdmap users used the Crowdmap platform was to create a map with a specific purpose or event (40%). About 63% of these users launched a map to cover an event in the city in which they live.

About 64% of Crowdmap users created a map for non-direct use, i.e., for training or demonstration purposes or simple curiosity. This in large part explains the power law distribution reported above. That said, about 30% found the technology counter-intuitive and too time consuming.

In terms of users who did launch a Crowdmap, about 16% felt that they were not able to generate the required public awareness vis-a-vis their map to make it as effective as they had hoped. In addition, around 19% noted they were able to make their map almost as effective as they hoped. However, 31% replied that they simply were not able to garner sufficient interest in their map. This finding also explains the power law distribution described above. Of those users who responded to the survey and felt they had been successful in raising awareness of their maps, about 23% noted that they had done so via engagement with community organizations, civic groups and/or professional organizations. Around 20% said they had successfully built public awareness by engaging with social media, SMS, etc. Only 6% of users said they used traditional news media to build awareness of their project.

A separate survey for semi-structured interviews was developed for users who launched high-profile projects using the self-hosted Ushahidi platform. A total of 37 high-profile projects were identified for the survey and seven respondents completed the survey, i.e., ~19% response rate. All seven respondents represented formal organizations and had used the platform in response to a complex humanitarian emergency or "natural" disaster. Two deployments were in developing countries and the remainder in the "Global South".

Some of the biggest challenges cited by users of the Ushahidi platform included "keeping reports up to date, embedding pictures and documents;" "the need to display the data on something other than a map;" "getting the word out quickly after

the launch of the site;" and "getting people to submit reports." When asked about failures, respondents' answers ranged quite widely, from persistent technical problems to mobilizing volunteer involvement. Many noted that using the platform was simply too time-consuming.

The findings from the quantitative analysis and surveys provide the first evidence-based analysis of crowdsourced data of its kind. In addition, the results supply actionable feedback to Ushahidi software developers on what they can do to improve their platforms and substantially increase the number of Crowdmaps that gain more traction and possibly greater impact. It should be noted that since this research over half-a-year ago Ushahidi Inc. has already been implementing a number of important changes including a set-up wizard, a wiki for Ushahidi users, and a review of the Crowdmap. In sum, this report provides an important baseline study—and indeed the only one of its kind—which could serve as an important comparison if this research is replicated in the coming years.

As we assess the growth and impact of Ushahidi in general and crowdsourcing in particular we should keep in mind that we are still at the very start of a transformative process. There is plenty more to do and learn before we can draw any firm conclusions, particularly vis-à-vis impact. Crowdmap, for example, is barely a year-and-a-half old, which means that users are still very much in the pioneering and discovery phase. Recall Clay Shirky's point that "technology only becomes socially interesting when it becomes technologically boring." This explains why the CrowdGlobe Project is intended to launch the means of an ongoing assessment of where we are now and what we can expect in the future—hence the interactive CrowdGlobe.net portal. This is not the final statement about crowdsourcing and Ushahidi. It is the opening statement of a new field of inquiry and civic action.

INTRODUCTION

Crowdsourcing is changing entire industries across multiple sectors. Coined by Jeff Howe in 2006 to describe businesses that were openly outsourcing small, incremental tasks to the general public, crowdsourcing has since become a major business sector itself. The multi-million dollar company Crowdflower, for example, crowdsources millions of tasks a year on behalf of top Fortune 500 companies.

Crowdsourcing has also had a profound impact on the not-for-profit sector. Wikipedia, of course, remains one of the most astounding examples of crowdsourcing to date. Even the venerable Encyclopedia Britannica announced that it would cease publication in the face of crowdsourced information platforms, especially Wikipedia. Crowdsourcing is also radically reshaping humanitarian response—a significant shift that many humanitarian organizations have yet to realize or fully understand, let alone respond to. While these humanitarian professionals were, for many decades, most often confronted by an information vacuum following a crisis, which meant that they were tasked with providing initial assessments, they are now confronted with a deluge of multi-media, user-generated content shared on multiple social media channels, often in real-time.

Disaster-affected communities are increasingly becoming digital. Thanks to the incredibly rapid commercialization of mobile phones worldwide, these communities have become the source of “Big Data” generated during the immediate aftermath of a crisis. Crowdsourcing is also disrupting the mainstream media industry as ordinary citizens are increasingly digital and this has catalyzed the global rise of citizen journalists. During the Arab Spring, for example, well over two-thirds of the video footage aired by Al-Jazeera was crowdsourced.

In sum, the majority of digital content shared online and via mobile phones is now user-generated, rather than produced by experts tied to formal institutions. This trend is not about to decelerate any time soon. Quite the contrary, the amount of user-generated, crowdsourced information will continue to increase exponentially.

This massive shift presents both significant challenges and important opportunities. Yet rigorous, data-driven research necessary to shed insights on this revolution in information

At a more theoretical level, CrowdGlobe offers an almost unparalleled opportunity to investigate the dynamics of digital information in political and policy processes.

is lacking. The purpose of Internews’ new CrowdGlobe initiative is to encourage and facilitate empirical research on the nature and impact of crowdsourced data—particularly geo-referenced data. CrowdGlobe is a platform agnostic, applied-research program that seeks to identify trends in both the use of crowd sourcing technologies and the data patterns generated by these new technologies. CrowdGlobe aims to produce in-depth reports on these trends and patterns. The project is accompanied by the interactive CrowdGlobe.net website, which provides access to underlying datasets, case studies and further analysis. In this way, the project seeks to catalyze additional user-generated analysis of crowdsourcing trends and patterns. In other words, the portal provides researchers with access to the meta-level crowdsourced data they need to understand both the opportunities and limitations created by new information and communication technologies. Through Crowd Globe.net, a global community of scholars and activists can now work together to create best practices in the context of sound analytics. At a more theoretical level, CrowdGlobe offers an almost unparalleled opportunity to investigate the dynamics of digital information in political and policy processes.

For the inaugural launch of the CrowdGlobe project, Internews elected to use Ushahidi as its first case study. This decision was



based primarily on ease of access to Ushahidi's crowdsourced data—particularly the hosted Crowdmapper data. In addition, at the time this research began in October 2011, some 12,000+ Crowdmappers had been launched in over 100 countries, which provided a size-able amount of data to analyze. Furthermore, the Ushahidi platform has been used for multiple purposes ranging from disaster response and human rights monitoring, to citizen journalism and election observation. Moreover, the Ushahidi team had long expressed a strong interest in deepening their understanding of how users were engaging their platforms and what the data they generated looked like. In particular, the Ushahidi team was interested in learning from this baseline study to both improve their technology and develop a real-time dashboard based on the metric identified in the analysis. To this end, Ushahidi generously provided all the meta-data used in this study and Internews partnered with faculty from George Washington University's (GWU) School of Media and Public Affairs (SMPA) to carry out a fully independent and rigorous analysis of this data.

It is worth repeating that the CrowdGlobe project itself is strictly platform agnostic and seeks to provide extensive assessments of all crowdsourced mapping platforms out there. Indeed, this first case study is simply a starting point to stimulate input from the wider crowdsourcing community and in the process identify additional case studies for future reports.

Ushahidi & Crowdmapper

See the Libya Case Study (page 23) or view online at <http://crowdglobe.net/deployments>

In 2008 when several bloggers and software developers responded to Ory Okolloh's call for some means to aggregate and

share the many reports of violence that were coming to her and other bloggers in Kenya, the world was less than a decade into the use of several establishing technologies for Ushahidi. Without them, Ushahidi would have been impossible.

Less than a decade before, for example, the remote sensing satellite industry had just gotten off the ground—literally—with the launch of Ikonos in 1999. It was the world's first high-resolution remote sensing satellite. It and a fleet of other satellites that followed helped create the highly detailed geographical information archives that makes open source digital mapping possible. Without georectified spatial data, that is, without data about a precise spatial reference point on a map, Ushahidi would not be possible. With georectified data, we are able to pinpoint locations and geotag them in relation to events that

By the year 2015, Sub-Saharan Africa will have more people with mobile network access than with access to electricity at home, some 138 million people.

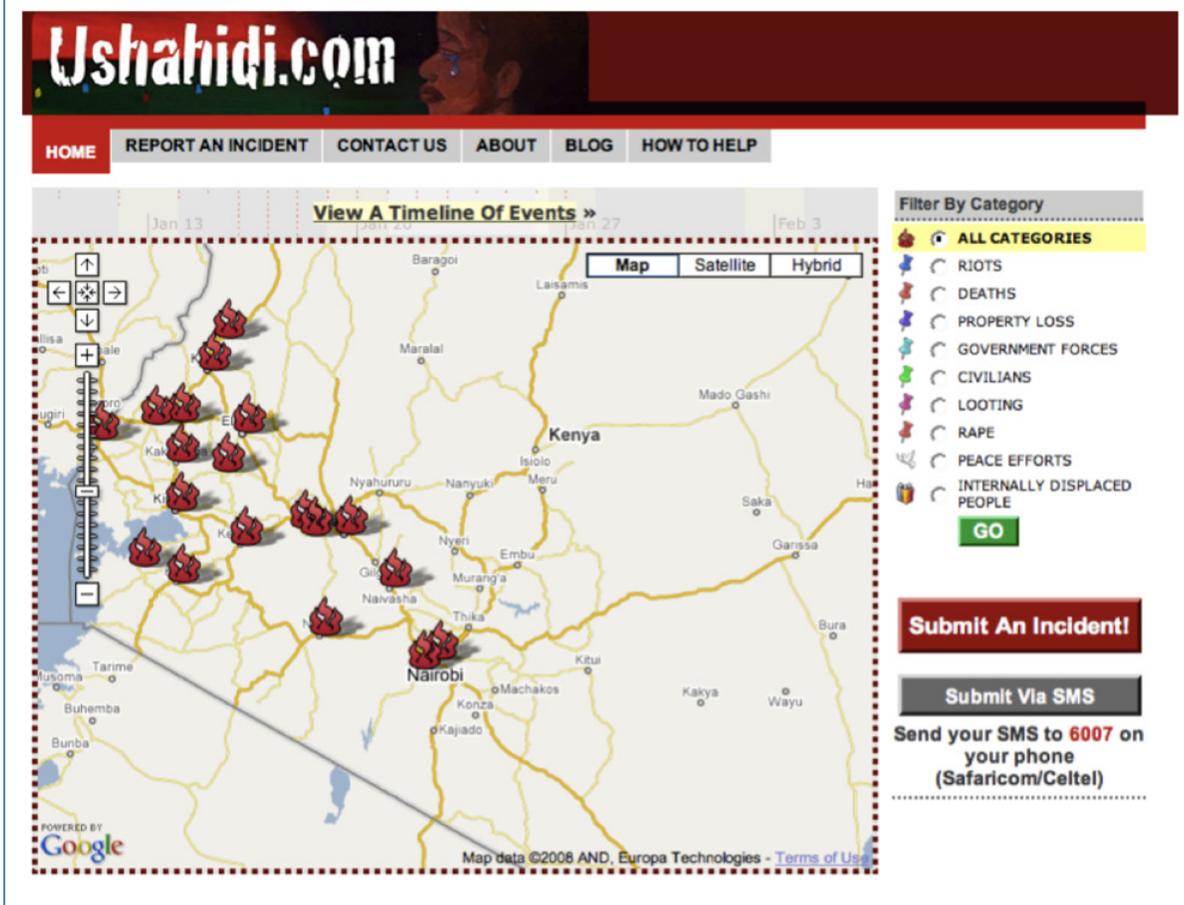
are significant in the context of a particular Crowdmapper deployment. Put another way, those little red dots would be meaningless and impossible without the relatively new capacity to use open-source geographical information systems that rest on remote sensing imagery. That system was eight years old when Ushahidi was first developed and Google Maps was only three years old.

Secondly, the remarkable growth of mobile telephony in the global south (and north) empowered publics (everyone with access to a handheld device) to be a potential part of a technologically enabled network that is global in scalability. In 2000, only 2 percent of the population of Africa had access to a mobile phone; in 2009, 28 percent did. By the year 2015, Sub-Saharan Africa will have more people with mobile network access than with access to electricity at home, some 138 million people. And by 2020 there will be at least one SIM card for every person on the continent. This same pattern is found all over the global south.

In Kenya the post-election violence of 2008 resulted in more than a thousand deaths and half-a-million displaced individuals. The Kenyan government tried to downplay the severity of the situation. At the same time, because journalists could not be everywhere and many human rights violations went unreported. Some Kenyan activists therefore decided to crowdsource and live map the crisis reporting. They set up a website with a Google Map of Kenya coupled with a web form and an SMS number. This allowed anyone with access to the Internet and/or a mobile phone to send in eyewitness accounts of human rights abuses. In this way, the "crowd" was able to document human rights abuses that would have otherwise gone undocumented.

The project was called Ushahidi, which in Swahili means, "witness" or "testimony". The Ushahidi platform is most simply described as a multimedia inbox connected to a live map. The inbox receives data from webforms, as well as emails, SMS, tweets, pictures, video footage, voicemails, etc. Users of Ushahidi can use various methodologies to collect this information—e.g., crowdsourcing—which they can subsequently

FIGURE 1: THE ORIGINAL USHAHIDI PLATFORM LAUNCHED IN 2008



map on a public website. The activists behind the Ushahidi subsequently launched a non-profit organization of the same name to provide the underlying technology as free and open source platform for others to customize and use for their own purposes. Since then, the platform has gone through several revisions and integrated additional technologies/media like smart phone apps and Facebook. In sum, the Ushahidi platform facilitates live, multimedia and collaborative mapping. Some 20,000 Ushahidi maps have been deployed in over 140 countries since the original Kenya Crisis Map in 2008.

In the Fall of 2010, the Ushahidi team launched Crowdmapping, basically a hosted version of the Ushahidi platform, which until then could only be used by downloading and installing the software on a computer. This often proved to be a challenging process, particularly for non-tech savvy users with little to no programming skills. Crowdmapping, on the other hand, is like "Google Doc"; no downloading or installing required. The launch of Crowdmapping has considerably lowered the barrier to entry

FIGURE 2: OF THE 20,000+ MAPS DEPLOYED SINCE 2008, OVER 15,000 ARE USING THE HOSTED CROWDMAPPING PLATFORM

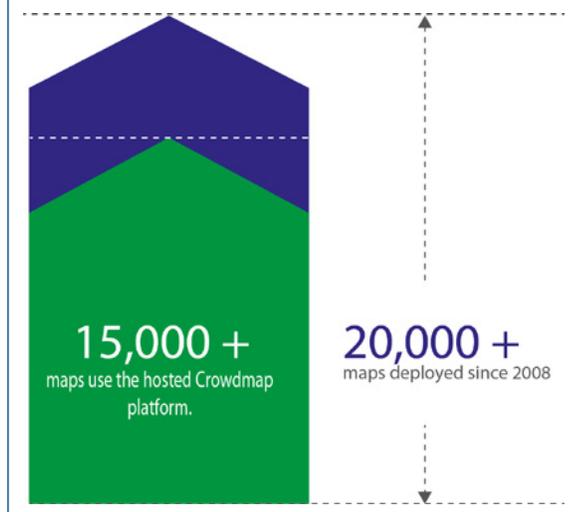
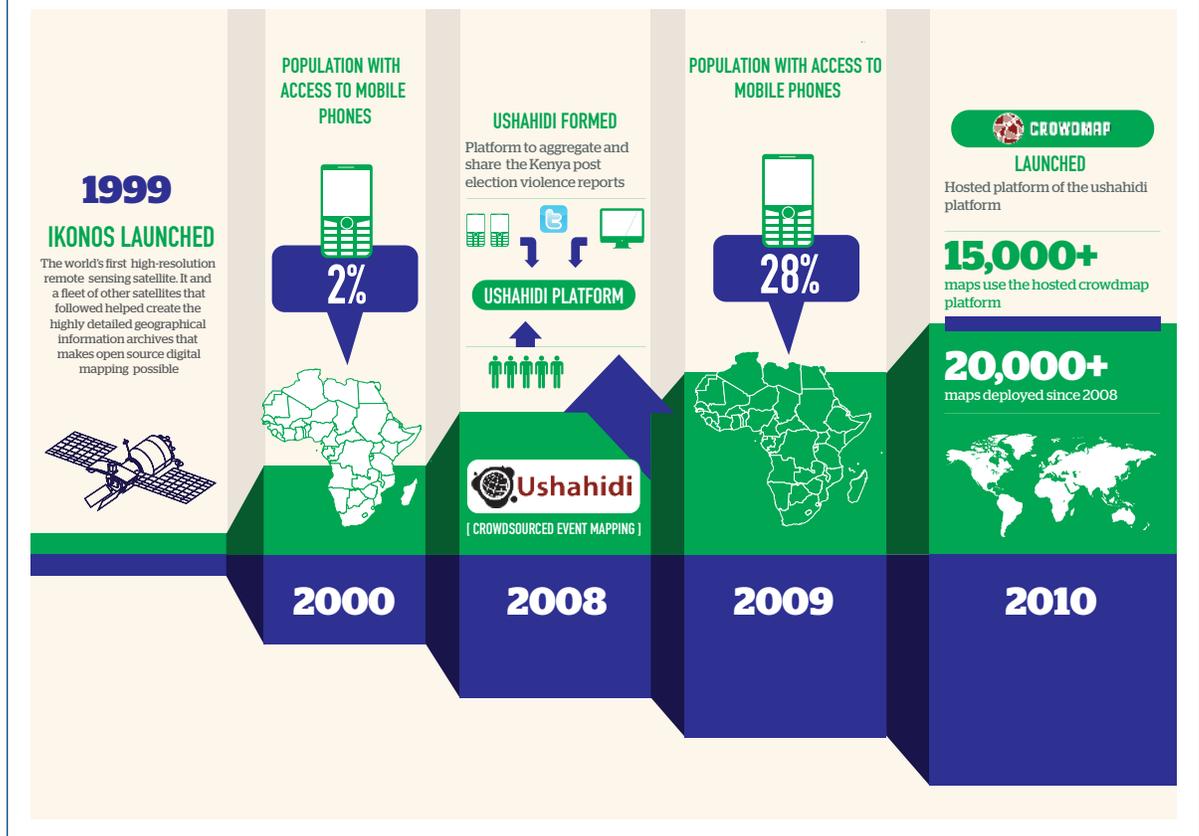


FIGURE 3: INFO-GRAPHIC ON THE TECHNOLOGY HISTORY LEADING TO THE CREATION OF CROWDMAP



for users interested in crowdsourced mapping. Indeed, of the 20,000+ maps deployed since 2008, over 15,000 are using the hosted Crowdmap platform.

The Ushahidi and Crowdmap platforms are not the only mapping technologies available for collecting and mapping crowdsourced data. The free and open-source OpenStreetMap platform, for example, is primarily known as a crowdsourcing project for street and road data around the world. However, the platform has also been used to map crowdsourced data related to other types of infrastructure such as refugee camps in Haiti, health facilities in Libya, damaged buildings in Turkey and disaster preparedness data in Indonesia. Google Maps has also been used to map crowdsourced data, as has the Google Map Maker platform, most notably in Southern Sudan. To date, however, the Ushahidi platform remains the only technology that is geared towards crowdsourced event mapping.

Research Questions And Design

When the Crowdglobe research project began in October 2011, 12,795 Crowdmaps had been launched by individuals and organizations around the world. Together with Internets, the George Washington University team and the authors of this report identified metrics that would provide insights on how users were using Crowdmap. They also identified metrics that could capture potential patterns generated by the Crowdmap meta-data. These metrics were broken down into three tiers.

The first tier included quantitative metrics that could be answered through a series of statistical queries. This included the topic/theme of the deployment; the number of reports per deployment; the number and type of categories per deployment; the number of users per deployment, etc.

The second tier required additional analytical work, drilling down into the context of the instances, which meant survey-based research. For example, what kinds of background do

TIER 1: Simple Queries over Raw Data	Tier 2: More Complex Queries Where Data Curation Is Needed	Tier 3: External Research And Analysis Required To Answer
<ul style="list-style-type: none"> - The name of Deployment - The date of first report - The date of last report - The default country (incl. lat/lon), or was this a multi-country deployment? - How many categories were set (at the time of our research)? - What are the names of the categories (at the time of our research)? - The total number of submitted reports - The total number of approved reports - The number of verified versus unverified reports - The count of reports from the following sources: Web, Email, SMS, Social media channels - The number of registered users (site administrators) - The number of aggregated RSS feeds (if any) - The number of unique visitors - The average number of reports per day - The number of multimedia contributions 	<ul style="list-style-type: none"> - The topic/theme of Crowdmap - The average time between submitted reports - The initial number of categories - The names of categories (one week after creation) - The names of categories (at the date of the last report) 	<ul style="list-style-type: none"> - What is the overlap of categories between similarly themed deployments? - What is the number of local vs. global contributions (project scope)? - Was Ushahidi, Inc. involved in the deployment? - What skillsets did the administrators/project leads have? - Who/what initiated the deployment? - Figuring the Gini coefficient (success/fail rates vs. access to tech, global volunteers, media attention, etc.)

Crowdmap users have and have they used this kind of technology before?

The third tier comprised metrics that could only be produced through direct engagement with those individuals who created online maps. What objective did users have when launching their crowdsourced map? Were they successful? This latter tier required individual semi-structured interviews.

In sum, the research did not focus on the highly engaged deployments or the deployments without engagement much beyond an initial investigation of the Crowdmap platform. Stated more precisely, we excluded from our analysis deployments of Crowdmaps with over 10,000 reports and those with fewer than 21 reports.

See the Haiti Case Study (page 21) or view online at <http://crowdlobe.net/deployments>

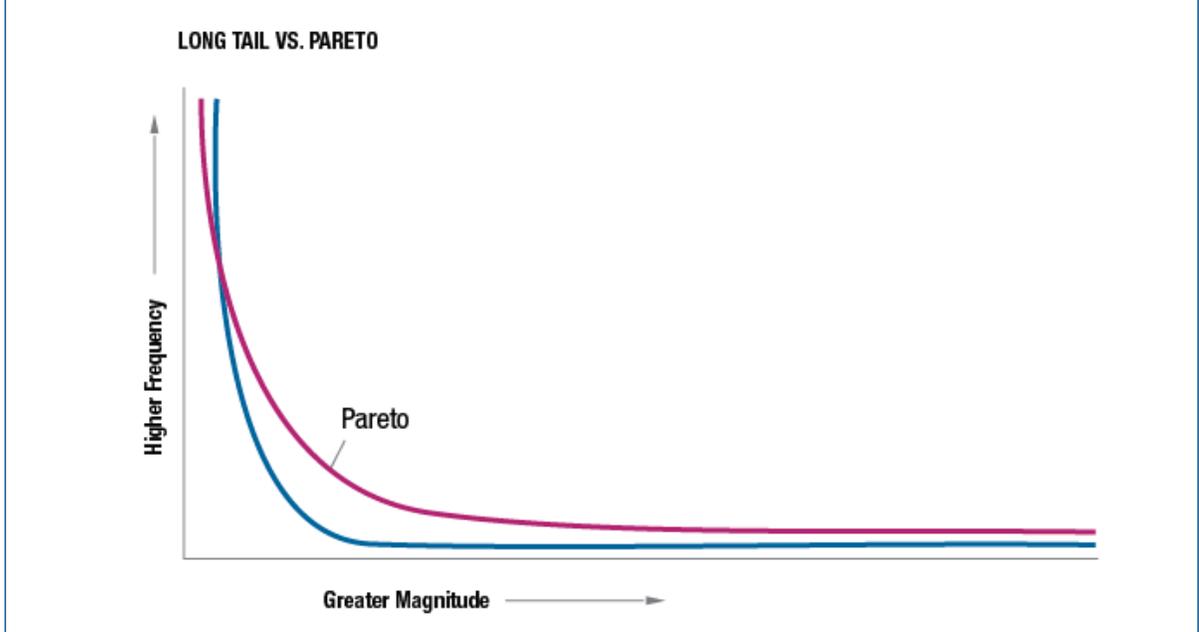
The GWU team's first major finding after analyzing the Crowdmap data for "Tier 1" metrics was the discovery of a "long tail" distribution.

In other words, the vast majority of Crowdmaps have very few to no reports while only a handful of Crowdmaps have thousands and even tens-of-thousands of reports.

Since Crowdmap is both free and relatively easy to use, this finding was actually not a surprise. With a few clicks of the mouse one can "deploy" a Crowdmap. Crowdmap is, by design, an easy access platform. With such a low barrier to entry, a power law distribution should be expected. It should also be kept in mind that more data does not necessarily equate with better or more successful uses of Crowdmap. The success of a given Crowdmap deployment is dependent on the purpose of the map, which is not always well defined, if it is defined at all. More data does not automatically imply greater impact.

Finally, focus on the long tail of deployments distracts us slightly from the impressive number of robust deployments that occurred between August 2010 when Crowdmap went live and October 2011 when the data were obtained. On average, there were 914 deployments per month. Future research will focus on the head of the long tail. So GWU researches conducted a "Tier 1" analysis on the "middle" of the tail, i.e., those Crowdmaps with reports numbering between 21 and 10,000. A total of 585 Crowdmaps (out of 12,795) fit this category, which was a more manageable number to work with vis-à-vis producing the "Tier 1" metrics.

FIGURE 4: THE LONG TAIL DISTRIBUTION VS THE PARETO DISTRIBUTION



For Tier 2 metrics, the team developed a survey, which was sent to all Crowdmapp users. The purpose of this part of our research was to learn more about Crowdmapp deployments from deployment administrators. What works well about Crowdmapp as a crowdsourcing platform and what doesn't? In terms of Tier 3 metrics, the GWU team engaged users of the downloadable Ushahidi platform. The GWU team wanted to know more about these more technically savvy users.

To this end, the team developed a separate set of questions for the purposes of carrying out dedicated semi-structured interviews. A total of five high-profile Ushahidi deployments were

selected: Haiti, Libya, Japan, Sudan and Egypt. In addition to the interviews, background research on these five case studies was carried out to provide more context to the analysis.

Quantitative Results & Analysis

"Tier 1" metrics were produced by running a series of database queries on each of the 12,795 Crowdmapps. The findings were then compiled into a single database, the main results of which are showed here:

FIGURE 5: RESULTS OF "TIER 1" ANALYSIS DONE ON THE ROW DATA OF THE CROWDMAPP DEPLOYMENTS

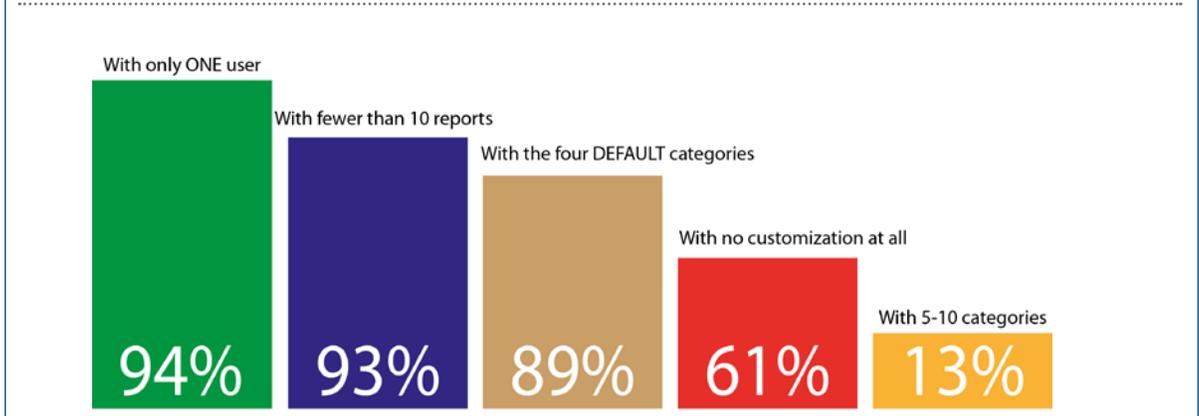
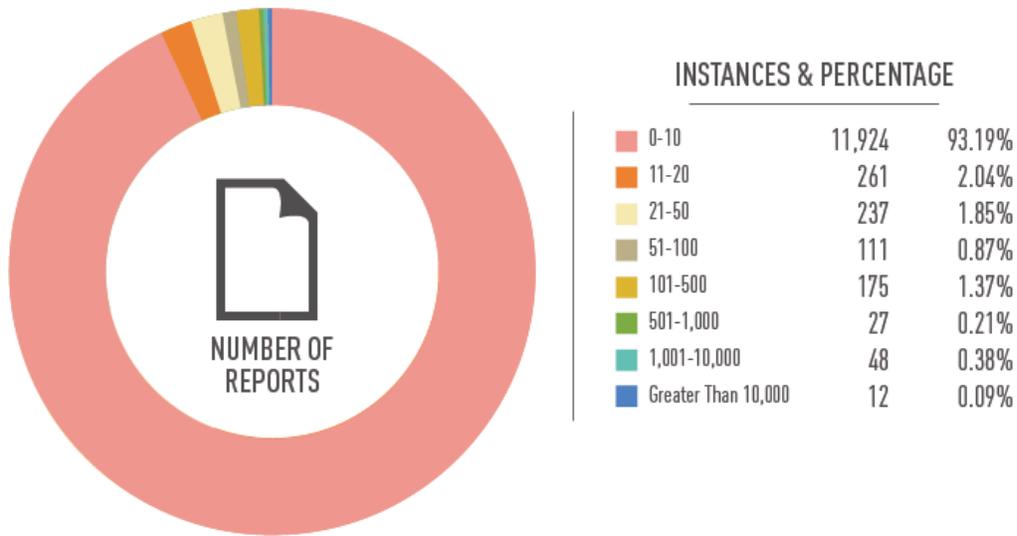


FIGURE 6: OUR INITIAL PROCESSING SHOWS THAT THE VAST MAJORITY OF DEPLOYMENTS HAVE LITTLE TO NO ACTIONABLE



The main results from these initial findings pointed to the next most logical step in our analysis. More focus was given to survey responses from both Crowdmapper users and Ushahidi users. Extant research suggested that the “Tier 1” analysis would reveal that an 80/20 “Pareto Principle” would govern the results.

The Pareto Principle is a name for a type of power law distribution: In the Crowdmapper case it would predict that 80% of Crowdmapper deployments would have little-to-no data, while the remaining 20% would have the vast majority of the data. Yet, as noted above, while about 61% percent exhibited virtually no activity beyond installation, 93% of Crowdmapper instances reported fewer than 10 reports. In short, the power law distribution was far steeper than the Pareto Principle would anticipate.

Power Law Distribution

This result prompted a question about the broader Crowdmapper ecosystem. Is this sort of distribution common across crowd-sourced phenomena? Many natural and social phenomena cluster around a mean or typical value. This distribution is captured by references to averages: average shirt size or average speed of cars on a freeway. The past decade and a half has seen an explosion of scholarly and popular interest in a different pattern—power law distributions.

Power law data have an inverse, exponential relationship between the magnitude of an observation and its relative frequency. Steep power laws combine concentration at the head with long, heavy tails. They have no “typical” value, as almost all of the observations are below the global average.

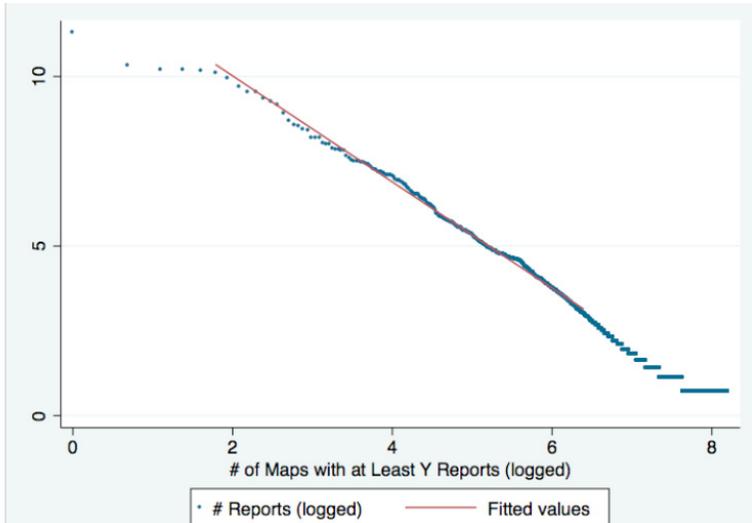
The size of earthquakes versus their frequency, the size versus frequency of solar flares, craters on the moon, wars, and even word use all follow this pattern. In English, for example, occurrences of *a*, *and*, and *but* are frequent (head), whereas words such as *oxymoron*, and *polymorphous* are used infrequently (tail).

The online environment seems particularly prone to the production of power laws in a host of different areas, from the ‘link’ topology of the Web to the size distribution of open source software projects, from blog traffic to the popularity of YouTube videos. Does the Crowdmapper data display the same empirical regularities?

Analysis of the Crowdmapper data suggests that the size of projects, as measured by reports, does indeed follow a rough power law. Formally, we would consider the Crowdmapper data to be power law distributed if the probability that a randomly selected map had *K* reports is proportional to $K^{-\alpha}$.

The figure below plots the size distribution of Crowdmapper projects on a log scale. On the Y-axis is the size of a project, as

FIGURE 7: A FIRST-DEGREE APPROXIMATION OF THE LOGGED NUMBER OF REPORTS VERSUS THE LOGGED NUMBER OF MAPS



A first-degree approximation of the logged number of reports versus the logged number of maps shows a power law distribution of “actionable” reports, as determined by the research staff: reports with more than 20 reports but fewer than 10,000.

measured by the (logged) number of reports on the map. On the X-axis is the (logged) number of maps that have at least Y number of reports.

The signature of a power law distribution is that it should form a straight line on a log-log scale. To a first-degree approximation, that is what we see here. In empirical data, it is common to observe that the very largest observations are smaller than a pure power law would suggest. This clearly seems true in the Crowdmapper data. A linear fit is overlaid on top of the empirical data; it covers deployments with at least 20 reports, and excludes the five largest map projects.

The body of the data is highly log-linear. Using the methods laid out in Clauset, Shalizi, and Newman (2007), we fit a power law to the data using maximum likelihood. These methods suggest a relatively steep value for alpha of 1.58. Deciding whether the Crowdmapper data most closely follow a power law or another roughly log-linear distribution (such as an extreme lognormal or a power law with an exponential cutoff) is not explored here, as this technical question makes no difference to the substantive conclusions.

Power law or log-linear distributions are found in so many domains, in part, because they can be generated by a host of

different underlying mechanisms. Merely observing a power law in the Crowdmapper data still leaves open the question of what underlying phenomena actually creates and sustains this pattern. One possible explanation for the log-linear pattern in the size of map projects is that they are, in fact, reflective of a power law distribution in offline data.

It is well documented that several of the phenomena that Ushahidi attempts to map—such as the magnitude of earthquakes or the size of armed conflicts—follow a power law. A handful of earthquakes or wars are massive, while most are small. If Crowdmapper contained data on every single damaging earthquake, for example, it is likely that the distribution of reports of damage or loss of life across maps would end up following a power law distribution.

Still, there is suggestive evidence in this data that there may be a compounding effect of success. The more reports a Crowdmapper project has, the more reports it seems to attract, leading it to a positive feedback loop. In physics, power law relationships often reflect phase transitions. It is possible that there is an analogous process by which a map project reaches critical mass. If confirmed, this may indicate the importance of strategies to get nascent map projects “over the hump.” This is a promising area for future research.

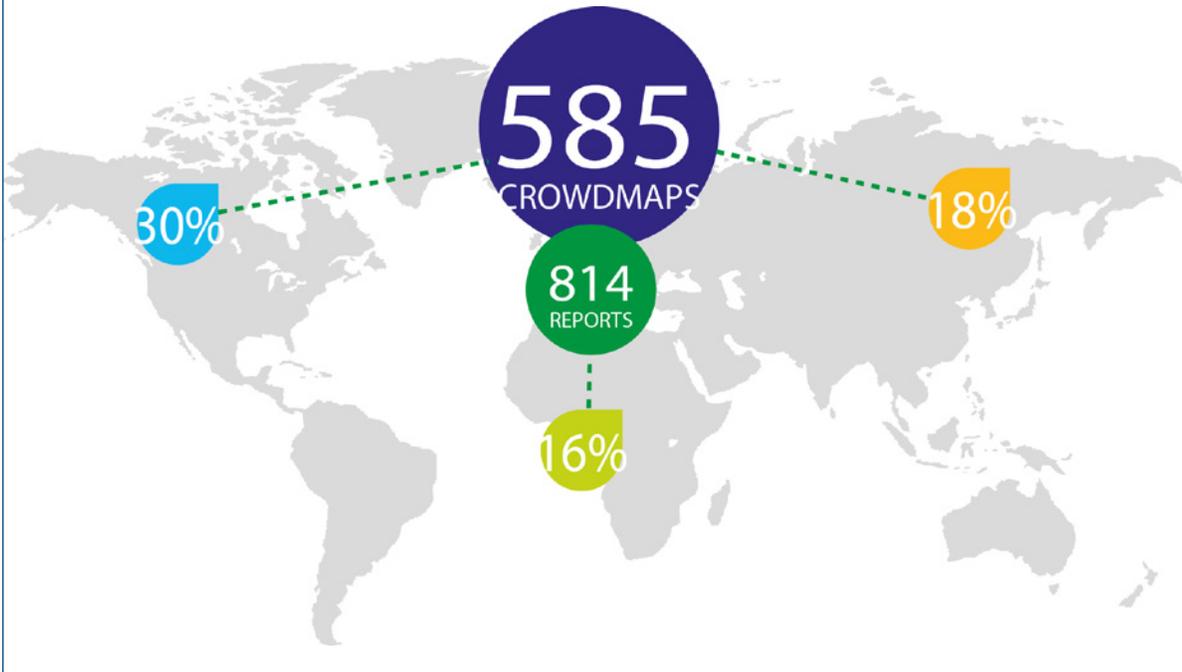
What distinguishes Crowdmapper users who appear to be mere “tire-kickers—analogue to those who go to an auto showroom only to look at, rather than buy, a car—from those who are fully engaged with the process? Since addressing this question would benefit Ushahidi with a much greater understanding of the strengths and weakness of their crowdsourcing platform, more attention was given to analyzing the 585 Crowdmappers that had between 21 and 10,000 reports.

A rigorous, systematic content analysis of the 585 Crowdmappers was carried out by two GWU graduate students. A GWU researcher designed the coding instrument and trained the coders. The results revealed that the vast majority of these (30%) focused on North America while 18% focused on Western Europe and 16% on Africa. On average, these Crowdmappers had 814 reports.

The median number of reports for this set of deployments was substantially lower, at 94, which is not surprising considering that the distribution of this set of cases is highly right-skewed. The content analysis also revealed what the Crowdmappers were being used for (see Appendix I for a dictionary for Crowdmapper data for the definitions of categories used below). Indeed, the

See the Japan Case Study (page 24) or view online at <http://crowdglobe.net/deployments>

FIGURE 8: GEOGRAPHICAL DISTRIBUTION OF THE CROWDMAPS ANALYZED



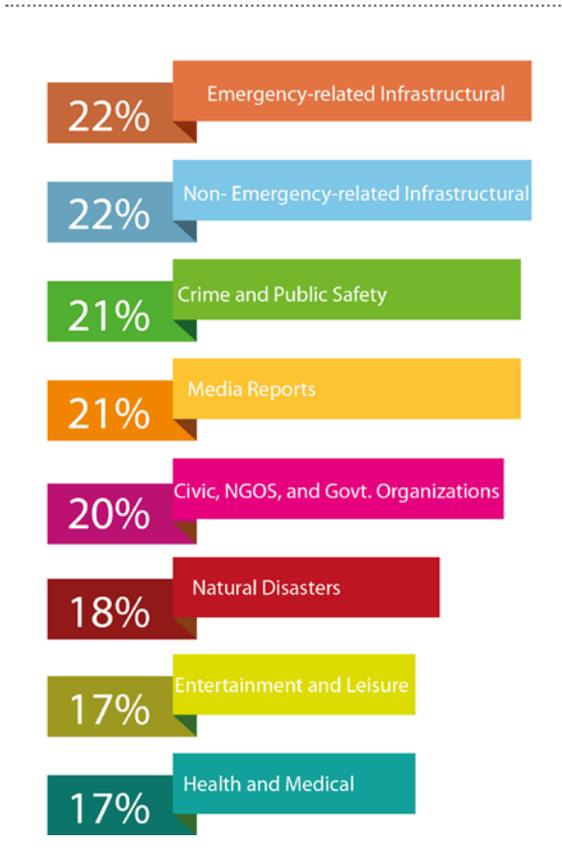
four most frequent themes that emerged across the data (see Fig. 2) include: Emergency-related infrastructural issues (which appeared in 22% of Crowdmap), non-emergency infrastructural issues (also 22%), crime and public safety issues (21%), and media reports (21%, which includes both traditional and new media outlets). The next most frequent set of themes included: Civic, non-governmental, and government organizations (20%), natural disasters (18%), entertainment and leisure (17%), and Health and Medical-Related Issues (17%).

Not surprisingly, however, when the analysis is broken down by region, the relative frequency with which these themes emerged in the regional deployments differs dramatically. For example, taking into account recent events, it is not surprising that the most common themes that emerged from the 63 deployments in the Middle East and Northern Africa pertain to: Crime and public safety issues (43%), human rights abuses (40%), emergency-related infrastructural issues (30%), and political organization (25%).

The distribution of themes in the 79 Western European deployments, on the other hand, paints a very different picture, with entertainment and leisure appearing in 32% of the deployments, followed by non-emergency infrastructural issues (25%), and media reports (23%).

It is also not surprising that the 16 deployments from the Caribbean region, 12 of which hailed from Haiti specifically,

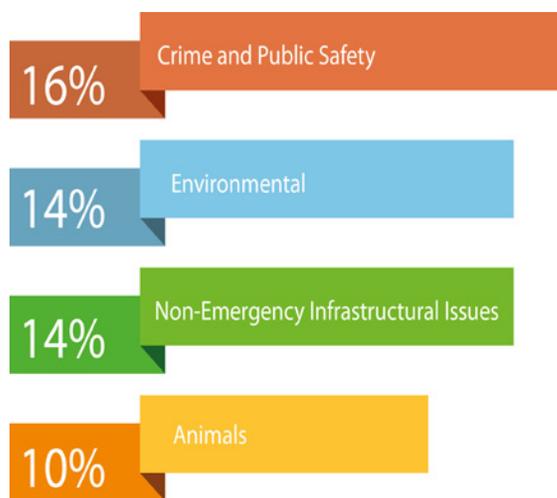
FIGURE 9: CROWDMAP MOST FREQUENT THEMES



heavily featured issues related to the occurrence and aftermath of a natural disaster (63% and 50%, respectively), emergency-related infrastructural issues (63%), health and medical-related issues (50%), and crime and public safety issues (38%).

As an additional point of comparison, it is interesting to note that there is limited difference in the frequency of specific themes when comparing the ends of the distribution in terms of number of reports. For example, comparing the themes that emerged in deployments in the lowest decile in terms of number of reports (i.e. those in the bottom 10 percentiles, with less than 32 reports) to deployments in the top decile (i.e. those in the top 10 percentiles, with more than 1053 reports), the majority of the themes appeared with commensurate propensity in both sets of deployments. In other words, the theme of crime and public safety issues was equally likely to appear in deployments with a limited number of reports (16%) as it was to appear in deployments with a large number of reports (16%). Among the handful that exhibited marginal differences, only three themes surpassed a differential of 10%-points or more. Specifically, these themes included issues pertaining to animals, fish, and birds (which were 10-percentage points more likely to be featured in deployments in the top decile), Environmental issues (which were 14-percentage points more likely to appear in deployments in the bottom decile), and non-emergency infrastructural issues (which were 14-percentage points more likely to appear in deployments in the bottom decile).

FIGURE 10: THIS GRAPH REPRESENTS THE PERCENTAGE OF MAJOR THEMES.



Regionally, deployments followed a logical pattern, often reflecting the primary concerns and needs that these communities faced. Where a country faced domestic stress, upheaval, or an exogenous shock, such as a natural disaster, this was clearly reflected in the deployments launched by citizens in those countries. However, it is also clear that, in regions characterized by greater stability, individuals employed crowdsource software to satisfy a diverse and broad range of interests and needs—from restaurant recommendations, to bird watchers documenting sightings of rare species, to mapping potholes and road repairs needed in a neighbourhood

Survey Research

The quantitative analysis has shed important insights on the patterns of Crowdmap users. In order to complement this

See the Sudan Case Study (page 25) or view online at <http://crowdglobe.net/deployments>

research and possibly explain the quantitative findings, this first CrowdGlobe report took a mixed methods approach, combining quantitative analysis with qualitative research. The latter forms the second part of the report below.

The purpose of the survey research was to complement the quantitative analysis carried out above. More specifically, the surveys are meant to place the numerical analysis into context and provide additional insights to the quantitative trends identified above. Since the quantitative analysis focused on both Crowdmap and self-hosted Ushahidi deployments, the research was composed of two surveys and case study analysis. For Crowdmap, a dedicated survey was sent out to 12,000+ users. The questions for this survey are listed online at crowdglobe.net/our-report/survey-data.

The Crowdmap survey was shared via Google Forms and 276 users responded to the survey, i.e., a response rate of ~2%. While this certainly does not constitute a random sample, the results are highly informative and relevant to the CrowdGlobe research project and Ushahidi.

A separate survey was developed for users who launched high-profile projects using the self-hosted Ushahidi platform. A total of 37 high-profile projects were identified. The second survey was thus shared with those individuals responsible for these projects. A copy of this survey is available on crowd globe.net/our-report. A total of seven respondents completed the survey, i.e., ~19% response rate. In addition to this survey, five case studies were selected for more in-depth, secondary research. The case studies, which included throughout this report are: Haiti Earthquake, Libya Crisis, Japan Tsunami, Sudan Election and Egypt Elections.

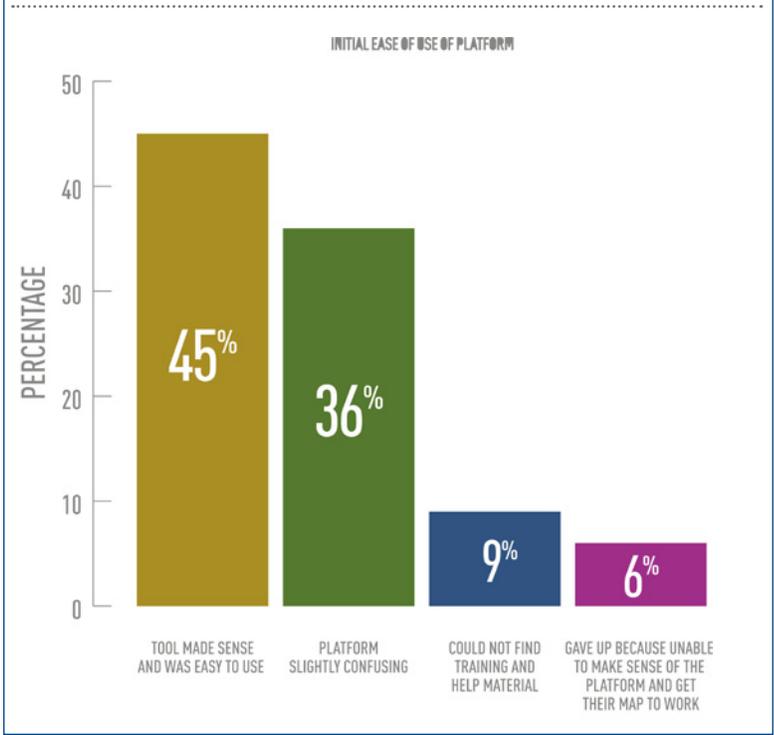
Survey Results & Analysis

The 276 surveys completed by Crowdmapper users provided interesting insights into the patterns of engagement with the Ushahidi software. Users' level of experience in using crowdsourcing & mapping technologies prior to trying out Crowdmapper was particularly limited. About 53% of users responded as having no prior experience, which is not surprising since the technology is still relatively new. About 27% noted that they had only used a similar technology once or twice before. About 80% of first-time Crowdmapper users were particularly new to this type of technology. Only 9% of users considered themselves as having considerable to extensive prior experience.

In terms of experience as a contributor to digital maps (as opposed to hosting), the distribution of responses was considerably less skewed. While 31% of Crowdmapper users had not contributed to digital maps in the past, about 22% of respondents noted that they had considerable to extensive prior experience in contributing to digital maps. About 30% of users answered that they had contributed to some digital maps in the past. As for prior experience in hosting a digital map, about 55% of users had never hosted one before while only 5% noted that they had a great deal of experience in hosting such maps.

The most important reason why Crowdmapper users used the Crowdmapper platform was to create a map with a specific purpose or event (40%). About 63% of these users launched a map to cover an event in the city in which they live. About 21% of users chose Crowdmapper for demonstration purposes while 35% used Crowdmapper to learn more about the technology and Ushahidi. In other words, more than half of Crowdmappers were created for non-direct use. Others wrote in specific answers such as "Great tool for training others on Ushahidi"; "To use it for my line of work"; "I work for UNHCR"; "All of the above"; "College project"; "Wanted to create a social mapping project"; "Wanted to test Ushahidi to make it better".

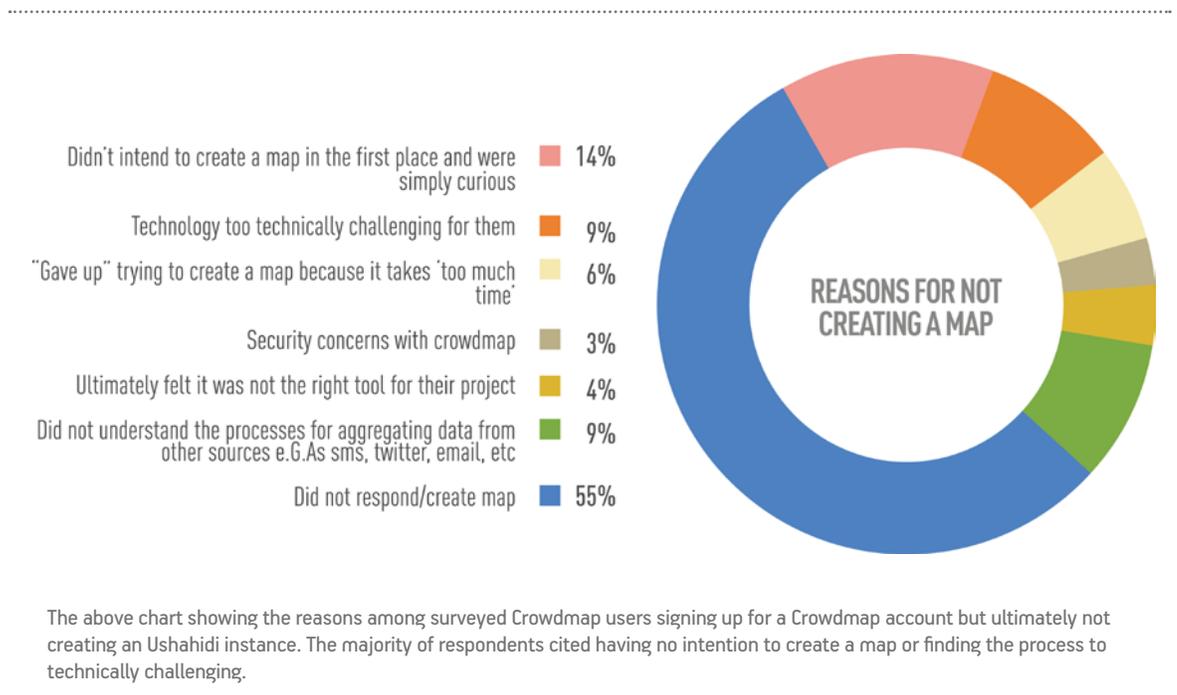
FIGURE 11: GRAPH OF THE INITIAL USE OF THE CROWDMAPPER PLATFORM



When users began to use Crowdmapper, about 45% of them found that the tool made sense and was easy to use while 36% of respondents explained that the platform was slightly confusing even though they were still able to figure it out and use the technology. Around 9% of users complained that they could not find training material or documentation to help them use the platform. Approximately 6% of respondents gave up because they could not make sense of the platform and were never able to get their map to work, which is a surprisingly low number.

As the quantitative analysis of the Crowdmapper meta-data revealed, the majority of Crowdmapper users who set up an account do not actually end up creating a map. Users cite several reasons for this. About 14% note that they didn't intend to create a map in the first place and were simply curious. Around 9% of users said the technology was simply too technically challenging for them to work with. Approximately 6% of users revealed that they "gave up" trying to create a map when they realized it would take too much time. About 3% of users had security concerns with Crowdmapper and therefore elected not to use the platform. Around 4% of users wanted to create a map but ultimately felt it was not the right tool for their project.

FIGURE 12: THE REASONS AMONG SURVEYED CROWDMAP USERS SIGNING UP FOR A CROWDMAP ACCOUNT BUT ULTIMATELY NOT CREATING AN USHAHIDI INSTANCE.



Lastly, approximately 9% of users did not understand the processes for aggregating data from other sources such as SMS, Twitter, email, etc.

Several users wrote in specific answers to explain why they never set up any map using the Crowdmap platform. These included "I was only using the platform as a demo"; "Others may not be able to easily navigate using the tool"; "Internal buy-off from organization created barriers—be nice to have quick summary of benefits/successes"; "the intended users did not accept it"; "was too confusing"; "I wanted to consult maps which were already done in the site...but I didn't find...or couldn't find"; "Wanted to create a map but the themes were too limiting"; "because I could not get data from expected sources"; "I wanted to show others the capability and determine if it would fit their needs"; "I would like to finish this, but I found it very challenging to complete the setup. Can you help me?" "The college project required deployment to azure—not enough time available to get PHP server running, tie in with sql server and get apache et al running on azure fabric"; "I only view maps"; "I needed more control for access. As such, I installed Ushahidi and have been using that"; "Mobile App needed improvement, need more controls for custom forms."

Of those users who concluded that Crowdmap was ultimately not the right tool for them, the most important reason cited

(by 18% of users) was that the platform could not be customized to meet their needs. About 7% of users replied that Crowdmap was too complex while 3% felt the tool was too simplistic. Others had more specific replies, such as "I did not understand what Crowdmap's capabilities were"; "need funding for project"; "cost of SMS messages"; "It is very hard to engage people in something new"; "data ownership"; "the context did not suit—low connectivity"; "I could not spend time learning the technology".

In terms of users who did launch a Crowdmap, about 16% felt that they were not able to generate the required public awareness vis-a-vis their map to make it as effective as they had hoped. Around 19% noted they were able to make their map almost as effective as they hoped while 31% replied that they simply were not able to garner sufficient interest in their map. Of those users who felt they had been successful in raising awareness of their maps, about 23% noted that they had done so via engagement with community organizations, civic groups and/or professional organizations. Around 20% said they had successfully built public awareness by engaging with social media, SMS, etc. Only 6% of users said they used traditional news

See the Egypt Case Study (page 27) or view online at <http://crowdlobe.net/deployments>

media to build awareness of their project.

About 30% of prospective re-users of Crowdfmap wrote that in the future they would consider using the platform if they identified a need. Around 10% would consider using Crowdfmap if they had more time. About 8% of users would use the platform if it was easier and/or technical support was offered.

Respondent demographics were also interesting. About 80% are men and the average age of a user is 40 years old. As for educational background, an impressive 43% of users have a post-graduate degree and a total of 84% have at least a college degree.

While only 19% of respondents completed the survey for the self-hosted Ushahidi deployments (i.e., not Crowdfmap), the results were nevertheless insightful. All applications of the Ushahidi platform were in response to a complex humanitarian emergency or “natural” disaster, with two deployments being in developing countries and the remainder in the “Global South”. The UN Office for the Coordination of Humanitarian Affairs (UN OCHA) accounts for the majority of deployments followed by media organizations (Washington Post and Australian Broadcasting Corporation). Only one deployment was carried out by a relatively small NGO. Most of the respondents indicated that they had minimal prior experience in using the Ushahidi platform before deploying their project. But most did note that they were already moderately experienced in contributing to digital maps.

Challenges

Some of the biggest challenges cited by users of the Ushahidi platform included “keeping reports up to date, embedding pictures and documents;” “the need to display the data on something other than a map;” “getting the word out quickly after the launch of the site;” and “getting people to submit reports.” Other challenges cited ranged from the difficulties in acquiring an SMS short code, being too dependent on the Ushahidi team to fix technical problems and bugs and information overflow. One recurring difficulty cited was fundraising to set up and maintain the project.

When asked about failures, respondents’ answers ranged quite widely, from persistent technical problems such as bugs to managing and mobilizing volunteer involvement. Many complained that using the platform was simply too time-consuming. Others, like humanitarian organizations, noted the “limited use of the system by traditional humanitarian entities,” which explained the “lack of understanding of how the system can be used,” and the fact that the system was just “too hard for high level decision makers to get what they needed.”

The media organizations that used the platform tended to highlight the interface as being problematic: “It was difficult to drill down and get information for a particular area or time period (e.g., last 24 hours) easily. There were also a number of features on the default interface that didn’t seem to work, such as the filter for reports with images and videos. “We wanted to try to customize the interface a bit more but there were limited options.” Furthermore, one media group added that their “journalists tried to verify some of the audience reports as they came in but found it too difficult and time consuming to do in reality.”

Successes

As for successes, respondents’ responses ranged widely as well. On the technical front, features such as RSS feeds, dynamic statistics and embedding mainstream news proved easy to use. Using the Ushahidi platform allowed one humanitarian organization to “mobilize key counterparts in the natural disaster monitoring and response field and learn about their interests and the potential uses they could give to the platform. It also allowed [them] to interact with local municipalities and departmental government in a way we had not done before.”

The fact that existing government data can also be mapped over time and space enabled one government to better understand the potential of crisis mapping. Another humanitarian organization documented how working with the technology and the Standby Volunteer Task Force (SBTF) has changed the way that they work internally, in terms of how they organize and process information. One major humanitarian group explained that their Ushahidi map “served to help stranded migrants and provide information on [the organization’s] operations. Another group explained that while the impact of their map was minimal, the project’s lessons learned were instrumental in launching their subsequent map, which provided an “alternative vision of disaster information that allows for various media to be brought in.” This in turn generated “better awareness for the responding organizations.”

One media organization revealed that their Ushahidi map received over 230,000 hits over a 3-week period and a total of 1,500 reports. This “showed the potential of real-time crowd-sourced mapping tools and showed that this sort of tool could be used in emergencies—at least in slow-moving ones such as floods—without too great an editorial risk.” Viewers were no longer passive observers but participants in the process. “It gave our audiences an opportunity to assist in reporting on an unfolding event as it was happening rather than ringing the local radio station they could directly submit a report to us

online. Ideally in the longer term we would like our news and radio presenters to be using the map as a reference tool as well as making call outs to the audience to submit reports." Another media group explained that citizens were able to leverage the Ushahidi map to organize their own response efforts, which "could not have been done without it [i.e., the map]." Finally, one NGO highlighted how the technology helped to inspire greater participation in a peace movement. This finding aligns with those identified in the short case studies on Haiti, Sudan, Egypt, Libya and Japan.

All respondents noted that they would use (or already had used) the Ushahidi platform again in the future. However one organization expressed some important hurdles: "We had some of our developers review the code for Ushahidi to consider further use and they reported that the code was not well structured or documented and would be difficult for us to build on and customize. There would also need to be improvements in the User Interface to make it more customizable."

CONCLUSIONS AND RECOMMENDATIONS

Out of the pain of the post-election violence in Kenya in 2008 emerged an ingenious tool for tapping into the potential power of people who, while physically separated, could become bounded together by electronic networks. As we've seen in this report, there are still encumbrances to the realization of this new potential. Some view the technology as dauntingly complex, or the social dynamics or public awareness necessities insurmountable. But other groups have tapped into this new powerful tool to create something entirely new in collective action dynamics

What this report has tried to do is set a standard for rigorous, high quality data analysis of crowdsourced data. It has, to be frank, offered a few surprises along the way. Yet it is important to offer a bit of perspective on Ushahidi in particular and crowdsourcing in general.

As we assess the growth and impact of Ushahidi in general and crowdsourcing in particular we should keep in mind that we are still at the very start of a transformative process. This report might be thought of as offering the equivalent of an Apgar score for a newborn. (The Apgar score refers to the results of several simple tests devised to measure the health and viability of newborn children immediately after birth). There is plenty more to do and learn before we can draw any firm conclusions, particularly vis-à-vis impact. Crowdmap, for example, is barely a year-and-a-half old, which means that users are still very much in the pioneering and discovery phase. Recall Clay Shirky's point that "technology only becomes socially interesting when it becomes technologically boring." This explains why the CrowdGlobe Project is intended to launch the means of an ongoing assessment of where we are now and what we can expect in the future. Hence the CrowdGlobe.net portal. To be sure, this is not the final statement about crowdsourcing and Ushahidi. It is the opening statement of a new field of inquiry and civic action.

Crowdsourcing is an important new tool of accountability. From its inception, Ushahidi has allowed motivated populations to hold governments accountable for misdeeds and mismanagement. From Kenya to Russia, crowdsourced information has provided insight into situations that otherwise would have remained out of sight and out of mind (at least for those not

caught up in the chaos). Because journalists and human rights or aid workers cannot be everywhere to monitor human rights abuses or the condition of desperate people caught up in desperate situations, crowdsourcing could bring a new level of awareness to circumstances such as these. The data we have presented show that this clearly does not always happen, and for a variety of reasons spelled out in our survey data. But the fact it has happened despite some of the stated challenges and that it may happen again may give pause to those in power who expect a free hand just because they have muzzled the press and intimidated some into silence.

On the research front, this initial study of Crowdmap data points to the necessity of careful empirical analysis done by those who are trained in research methods appropriate to the analysis of network dynamics. In particular and as noted above, our discovery of the prominence of the long tail in Crowdmap deployments led to a greater research emphasis on understanding the factors that prevent curious "tire kickers" from becoming active users. Coupled with sophisticated quantitative methods is the need for in-depth qualitative field analyses of deployments. We have not been able to do that here, relying instead on the self-reporting via survey-based research. Future research should investigate the factors that lead to successful deployments, and those that do not. This is precisely why the CrowdGlobe.net website has been launched—to start a conversation and a collaborative effort.

This takes us to what is, in a sense, our most important preliminary conclusion. This project's greatest contribution might come in the form of establishing best practices for users and potential users, and perhaps even new procedures at Ushahidi.

For example, despite the surprisingly high levels of education among many respondents, many still reported that the complexity of the platform impeded their full use of it. This suggests that further refinement and greater ease of use would be beneficial.

The good news is that Ushahidi has already taken numerous steps to address these and other challenges identified in this report. Crowdmapper, for example, now includes a dedicated “wizard” to guide first time users through the customization process. The organization has also recruited a full time Community Manager who has organized more end-user meet-ups in 2012 than in all four previous years combined. In addition, the Community Manager has launched a dedicated wiki to provide significantly more documentation on how to use the platform. Crowdmapper developers are also developing a public online library to facilitate the discoverability of Crowdmappers

that will give them far more visibility. Indeed, as the survey results clearly showed, strong media and community outreach is critical to gaining traction. Furthermore, Ushahidi’s user-interface design team has also carried out a full usability review of the Crowdmapper platform, with substantial changes on the way. Finally, Ushahidi has partnered with TechChange to offer a dedicated, hands-on course on how to use Ushahidi/Crowdmapper and create successful deployments. These important improvements, taken together, are bound to generate Ushahidi/Crowdmappers that gain both more traction and visibility. In sum, this report provides an important baseline study—and indeed the only one of its kind—which could serve as an important comparison if this research is replicated in the coming years. In the meantime, we hope that further analyses will help us understand the power law dynamics of crowdsourcing, leading, eventually, to a less pronounced effect, if only at the margins.

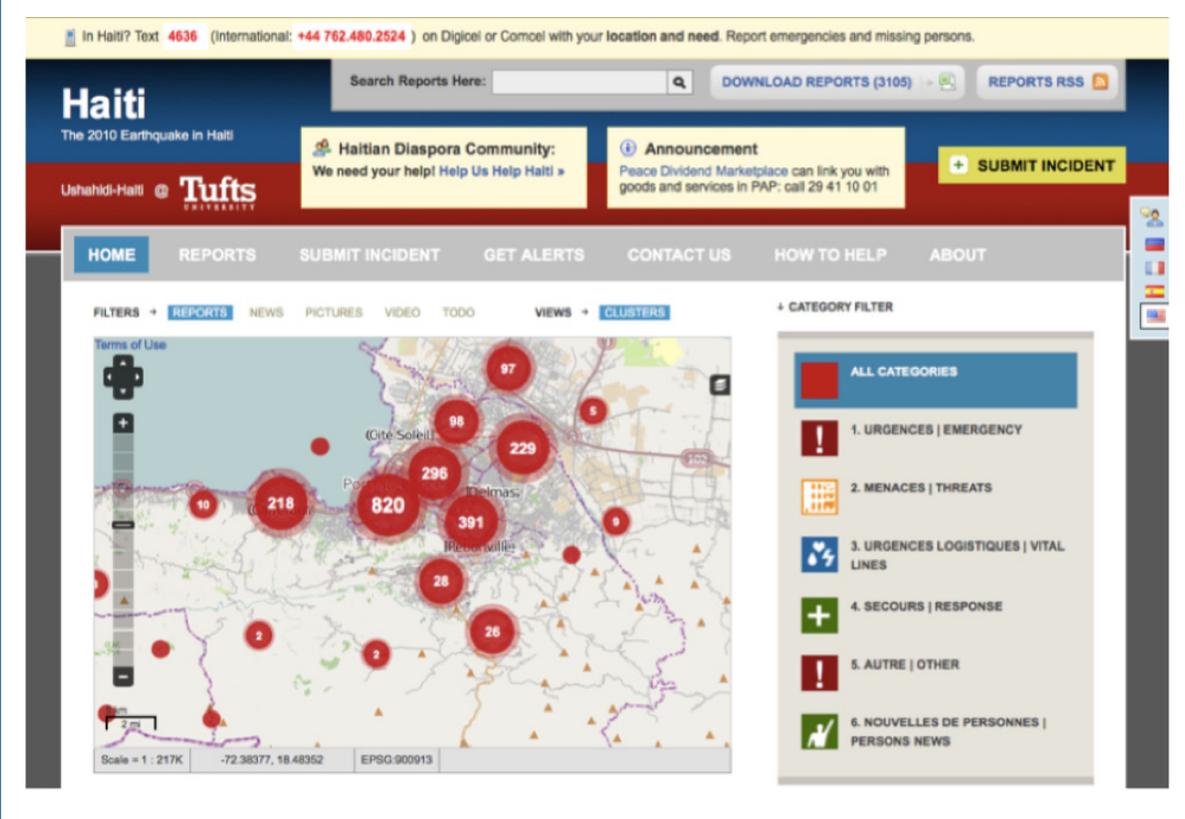
CASE STUDIES

CASE STUDY 1: Haiti

A devastating earthquake struck Haiti on January 12, 2010, resulting in hundreds of thousands of lives lost. Within hours, the Ushahidi's Patrick Meier and David Kobias launched a live Ushahidi Map of Haiti. During the first few days, the content mapped on the Ushahidi platform was sourced from Twitter, Facebook and other online sources such as mainstream media. Soon, they couldn't keep up with the deluge of information

on Haiti. So Patrick Meier reached out to colleagues at Tufts University for support and by the end of the first week had trained more than 100 volunteers on how to live map Haiti. During this time, a free SMS short code was secured from Digicel, Haiti's main telecommunications company. This allowed anyone in Haiti to text in their most urgent needs and location, which could then be mapped on the Ushahidi platform.

FIGURE 13: THE USHAHIDI-HAITI CRISIS MAP SPEARHEADED BY TUFTS UNIVERSITY



Some ten days after the Haiti map was launched, the head of the US Federal Emergency Management Association (FEMA), Craig Fugate, noted that the live map provided the most comprehensive and up-to-date information available to the humanitarian community. What is striking about this statement is that the map was not launched by FEMA, or the United Nations, or any professional humanitarian organization, for that matter. The live map was launched by student volunteers from a dorm room in snowy Boston some 1,500 miles away from Haiti. Over 3,000 reports were mapped and according to the Marine Corps and US Coast Guard, the Haiti Crisis Map helped them save hundreds of lives. But this live map would not have been possible, were it not for two other equally remarkable volunteer-led initiatives in Haiti: OpenStreetMap and Mission4636.

In the wake of the Haiti earthquake, the Google Map for the Port-au-Prince area was highly incomplete. This made it very difficult to find street names let alone specific addresses when mapping information on the Haiti map. The OpenStreetMap (OSM) community came to the rescue by crowd-sourcing the most detailed and comprehensive map of downtown Port-au-Prince. They did this by tracing satellite imagery and making their map openly and freely available. Some 600 volunteers from several dozen countries contributed over one million edits to the Haiti OSM map during this period. Needless to

say, the team at Tufts quickly switched from Google Maps to OpenStreetMap as a result.

Meanwhile, Mission4636 was in full swing. Thanks to Digicel's support, anyone in Haiti could text the number 4636 for free to communicate their most urgent needs and location. In this way, as in Kenya, the team leveraged the high prevalence of mobile phones to crowd-source needs assessments in real-time from the disaster affected communities in Port-au-Prince. The incoming text messages, however, were in Haitian Creole. They needed to be translated into English if the information was going to be used by the humanitarian community. This is where Mission4636 came in. Within a week, over 1,000 volunteers from more than 40 countries were recruited (primarily via Facebook) to translate incoming SMS's in near real-time. These volunteers were predominantly from the Haitian Diaspora community. Together, they translated some 80,000 text messages, with an average turn-around time of 10 minutes per SMS. Mission4636 volunteers also mapped these messages so they could be added to the Haiti map being curated by the volunteer team at Tufts University. Together, OSM-Haiti, Mission4636 and the Haiti Crisis Map demonstrate the important role that new crowdsourcing technology can play in supporting humanitarian response.

CASE STUDY 2: Libya

On March 1, 2011, the head of the Information Services Section (ISS) at the UN Office for the Coordination of Humanitarian Affairs (OCHA) reached out for help. He and his team needed to get a better picture of the humanitarian crisis unfolding in Libya in order to plan their response operations. OCHA had no personnel on the ground at the time and the ISS team could not rely on the information produced by the regime in Tripoli. But the Arab Spring had shown how much information existed in the social media space, and therefore ISS Head Brendan McDonald called for a live Crisis Map of Libya to better inform their situational awareness.

Within hours of this request, a live Crisis Map was launched. Operational crisis mapping in hostile environments presents some important challenges, not least of which is security. This explains why the two live maps were produced. One was password protected and exclusively for the humanitarian

community while the other was public but on a 24-hour time delay and with heavily redacted information.

Just days after the launch, the Executive Director of the World Food Program (WFP), Josette Sheeran, noted that the live map provided an excellent resource to plan her agency's relief operations along the Egyptian and Tunisian borders. Like all the other case studies featured above, the Libya Crisis Map was made possible thanks to a vast volunteer network. Unlike some of the earlier crisis mapping efforts, this network was organized and prepared. The Standby Volunteer Task Force for Live Mapping (SBTF) is a global volunteer network of some 700 volunteers based in close to 70 countries who are trained in live crisis mapping operations. Many of these volunteers are veterans from the Kenya, Haiti and Russia Crisis Maps. Together, they mapped over 2,000 reports on the Libya Crisis Map. So when the UN's Brendan McDonald reached out for help, he was actually requesting the activation of the SBTF.

FIGURE 14: THE LIBYA CRISIS MAP LAUNCHED FOR THE UNITED NATIONS



In McDonald's own words, the SBTF provided invaluable support to the UN's operations in Libya. In an thank-you email addressed to the volunteer network, the head of ISS wrote the following: "Your efforts at tackling a difficult problem have definitely reduced the information overload; sorting through the multitude of signals on the crisis is no easy task. The Task Force has given us an output that is manageable and digestible, which in turn contributes to better situational awareness and decision-making. I look forward to further developing our collaboration in this emergency and beyond."

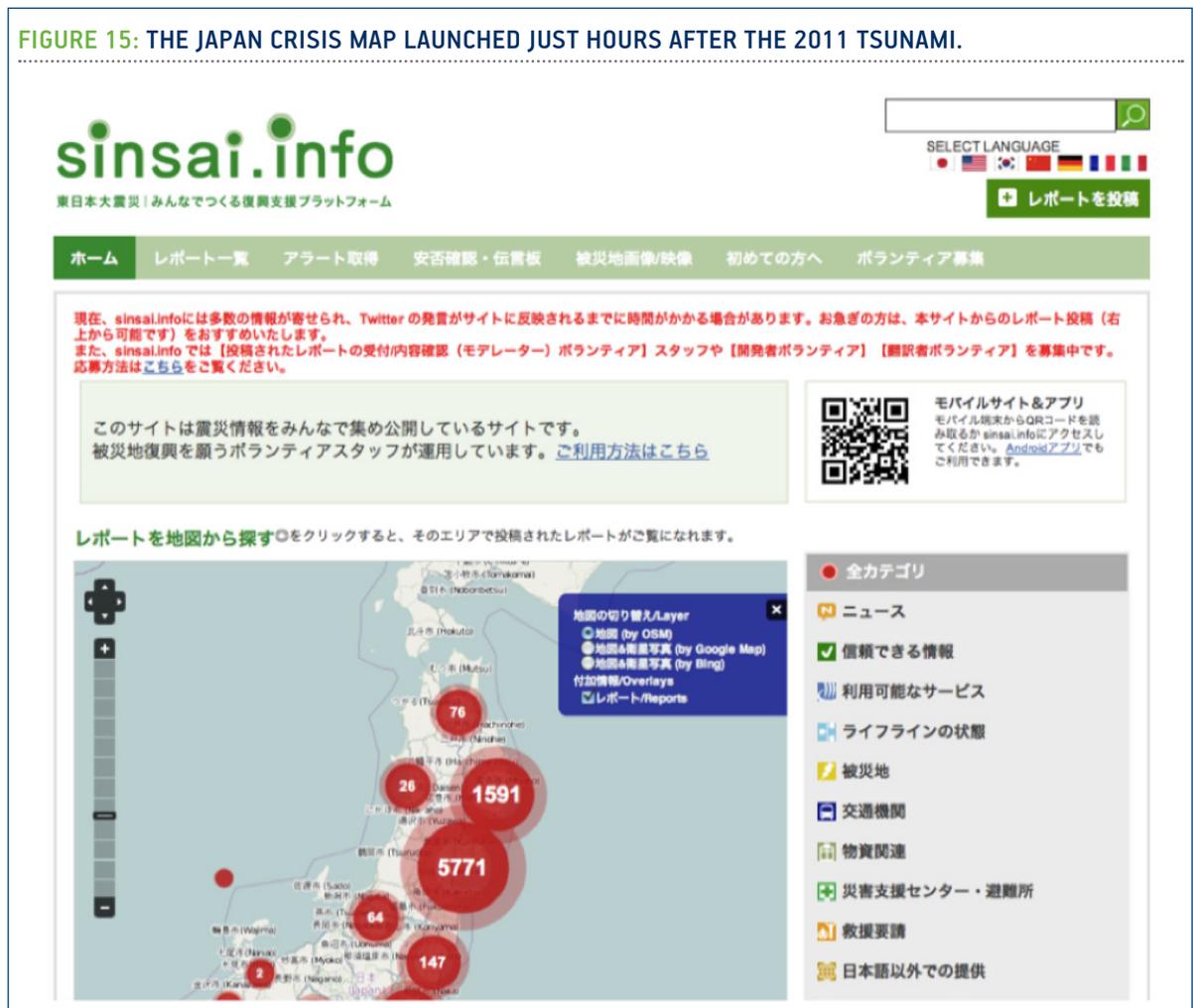
Around the same time that the Libya Crisis Map was launched, the International Organization for Migration (IOM) also launched a live crisis map to help inform their decision-making.[3] The purpose of this live map was to monitor the migration crisis resulting from the violent conflict in Libya, which resulted in the need to evacuate thousands of stranded migrants along the borders of Egypt, Tunisia and later Niger.

CASE STUDY 3: Japan

During the Libya Crisis Map deployment, another crisis on the other end of the planet struck. The devastating earthquake that shook northern Japan on March 11, 2011 resulted in a massive tsunami that took tens of thousands of lives and destroyed critical infrastructure such as mobile phone communications. Inspired by what they had seen in Haiti, volunteers from the Japanese OpenStreetMap (OSM) community launched their own live crisis map and shared the word via multiple social media channels

The OSM volunteers who were operating out of Tokyo closely monitored the Japanese Twittersphere, mapping an average of 3,000 tweets per week during the first month of operation. In this way, the Japan Crisis Map provided the most comprehensive and up-to-date information available on the impact and resulting needs. In total, the Sinsai.info map was accessed by over half-a-million users most of whom were based in the

FIGURE 15: THE JAPAN CRISIS MAP LAUNCHED JUST HOURS AFTER THE 2011 TSUNAMI.



disaster affected area. Obtaining relevant data in a timely manner from the Japanese government was difficult, which made the Sinsai.info deployment even more important. Again, however, ordinary volunteers spearheaded this initiative, and most of them had never done anything like this before. The author and other volunteers from the SBTf also provided the Japanese team with both strategic and technical support.

CASE STUDY 4: Sudan

Sudan Vote Monitor was the first time that an Ushahidi platform was deployed in a country under authoritarian rule. The pilot project was led by the Sudan Institute for Research and Policy (SIRP) and Asmaa Society for Development, in collaboration with other Sudanese civil society organizations. The purpose of deployment was to utilize the Ushahidi platform to support the independent monitoring and reporting of Sudan's first multi-party elections in 24 years.

The initiative complimented the paper-based independent monitoring efforts of formal election monitoring groups and offered Sudanese NGOs and the public at large an independent, online platform for election observation for the first time in Sudan's history.

The Ushahidi platform was considered particularly useful in Sudan, Africa's largest country, where long distances and inadequate infrastructure posed a significant challenge to civil society election monitors.

The spread of mobile communications throughout the country in recent years offered a unique and feasible opportunity to utilize SMS to overcome this challenge. Participating civil society groups deployed over 2,000 independent local observers throughout the 15 northern states. According to SIRP, "these observers continuously reported back what they witnessed at various polling stations across these states, using standard paper reporting forms. When texting, they used code, e.g., 1 = election fraud, 2 = voter intimidation, etc. This was done to

FIGURE 16: SCREENSHOT OF THE SUDAN VOTE MONITOR PLATFORM AFTER THE ELECTIONS



provide more cover to the citizen monitors." It is unclear how many text messages were received, however.

The site went live on April 10, 2010 with web and SMS reporting in both English and Arabic to coincide with the start of the elections held April 11-15, 2010. Response was relatively strong both inside and outside the country given that this was the first project of its kind in the Sudan. According to SIRP, "a total of 564 reports were received from the web (or translated from paper-based forms) from 419 locations, covering 26 election-monitoring categories. The web-based platform attracted wide interest from citizens, a variety of international organizations active in Sudan, as well as the local National Telecommunication Commission." However, the Sudan Vote Monitor website was blocked by the Sudanese government for two days before it was unblocked following US government pressure. In addition, fewer than 300 reports are actually mapped on the Ushahidi platform and certainly not from more than a dozen or so different locations.

In general, and compared to the other Ushahidi case studies figured in this report, the Sudan Vote Monitor project is not generally considered a success. The project was not well organized and only came together at the last minute. Indeed, the SMS short codes that were used for the project were only made available the night before the elections. So the organizers had had no time to get the word out about the SMS let alone carry out trainings or any simulations. This perhaps explains the fact that only 300 reports were submitted when there were a reported 2,000 election observers.

CASE STUDY 5: Egypt

In the Fall of 2010, the Egyptian and Cairo-based Development and Institutionalization Support Center (DISC), used the Ushahidi platform to launch U-Shahid; the goal of which was to monitor the parliamentary elections in November and December 2010. This independent initiative became particularly important when the Mubarak regime announced that it would not permit any official international election monitoring groups into the country.

The project was rather simple on paper: use the Ushahidi platform to monitor the elections by allowing people to send SMS, Tweets, Facebook comments, voice mail, e-mail and reports via web-form to the live map. DISC decided to draw on both crowdsourced reporting and "blogger-sourced" information. This meant getting the word out to the wider public while navigating the restrictions imposed by Egyptian national security, and also training a large network of trusted bloggers across

the country. Despite government restrictions, training for these bloggers took place in 5 major cities: Cairo, Alexandria, Assyut, Mansoura and Port Said.

On the technology side, DISC translated their Ushahidi platform entirely into Arabic since the U-Shahid project was not meant for an international audience but rather an Egyptian one: "an Egyptian project for Egyptians" noted one blogger. Egyptian software developers integrated Twitter, Flickr and YouTube with Ushahidi. Since Facebook was and continues to be an important platform for Egyptian youths, the group also created a Facebook feature that enable comments on a Facebook wall to be easily mapped on the Ushahidi platform.

During the elections, DISC mapped 2,700 reports, which included 211 supporting pictures and 323 videos. The team of Egyptian bloggers was also able to verify more than 90% of the content that ended up on the map by using basic journalist techniques such as triangulation and follow-up. Most of the mapped reports, however, came from the pre-established network of trusted bloggers, which did not require immediate verification. In total, the web-based map received close to 60,000 hits, the vast majority of which came from within Egypt. Interestingly, the next highest number originated from Saudia Arabia with just under 5,000 hits. The group was also proactive in disseminating this information, printing press releases and combining both new and traditional media for maximum impact. Their efforts were featured on Egyptian television, on BBC Arabic and dozens of articles in ten different languages. Indeed, both local and global media used the data generated by U-Shahid as part of their election coverage.

Naturally, the project also got the attention of the Egyptian government. Surprisingly, however, this attention began even before the project formally launched. The Egyptian state contacted DISC's director Kamal Nabil when the program design was still being developed. The government official told Nabil that his name was recurring "too often" in phone conversations between activists. The Egyptian Ministry of Interior subsequently shadowed the project in different ways: by tapping the cell phones of bloggers who comprised the core team; by requesting copies of the agendas for all meetings related to U-Shahid; and by requiring that a list of all individuals trained on the use of the platform be submitted to them. E-mail addresses, Facebook pages and Twitter accounts of the core team were reportedly all under surveillance since the start of the project, and the Ministry of Interior openly asked Nabil what his reaction would be if they were to shut down the U-Shahid project before the elections.

DISC was not immune to this government strategy: several

FIGURE 17: SCREENSHOT OF THE U-SHAHID PLATFORM AFTER THE ELECTIONS



new Facebook groups were launched to engage in personal attacks against the core team by accusing them of being affiliated with the United States, under the pretext that they had participated in a Freedom House-organized conference in DC earlier that year. Some of those Facebook groups called on young Egyptians to “watch out” for projects that could endanger the national integrity and the political independence of the country. Activists reacted to these attacks by conducting a virtual battle. Once a government-supported group was identified, dozens of activists would write on the group’s wall and basically occupying the entire wall with counter opinions. One of these Facebook groups was completely overrun after the group’s name was changed from “Youth for Funds” (sarcastic), to “State Security for Intimidation.”

DISC was well aware that technology alone would not change the political situation in Egypt. They also knew that Egypt’s National Security could shut down the project and block access to the website whenever they wanted. Furthermore, everyone involved in the project knew full well that their involvement in U-Shahid could get them arrested. As recent events have clearly shown, countries like Egypt and the Sudan are particularly agile in surveilling digital activists during election periods. But this did not discourage the Egyptian activists. The ability to do something different, to have an alternative was enough to be the difference. At the end of an U-Shahid training workshop in Cairo, one participant spoke with the lead trainer and simply said: “You know? We may all end up in jail, but before this I thought there was no hope to change anything. Now I can even dare to think it is worth a try.”

APPENDIX I: DICTIONARY FOR CROWDMAP DATA

Civic and Neighborhood Organizations

- Civic Organization
- Support Groups

Corruption (Non-election related)

- Bribery
- Corruption other

Crime and Public Safety (Non-War/Non-Protest related)

- Attack
- Gunshots
- Looting
- Murder/Deaths
- Presence of Military
- Presence of Police
- Rape
- Riots
- Sexual Assault
- Sexual Harassment
- Theft
- Vandalism
- Arson

Education and Schools

- Educational Programs and Training
- Schools

Elections

- Campaigning
- Election Results
- Elections Other
- Fraud (Clear instances of rigging and ballot tampering)
- Problems at Polls (e.g. Missing Ballots Voters, Names Missing, Problems with Machines, Long Lines, etc.)
- Riots/Protests by Voters
- Voter Intimidation

Emergency-Related Infrastructure Problems

- Donation/Fundraisers
- Electricity Outage
- Emergency Messages and Info
- Food Shortage/Needed
- Medical Care needed
- Relief/Aid Needed
- Shelter Provided
- Roads Damaged
- Shelter Needed
- Water System Problems
- Relief/Aid Provided
- Medical Care Provided
- Food Provided
- Water Provided

Entertainment and Socializing

- Entertainment
- Social Gatherings
- Where to Eat
- Shopping
- Hotels

Environmental Issues

- Animal/Fish Deaths
- Over-Grazing
- Pollution
- Environmental Issues Other

Human Rights Abuses

- Citizen Attacked/Beat by Police/Military
- Citizen Killed
- Disappearance/Kidnapping
- Excessive and Inhumane Punishment
- Forced Displacement
- Suppression of Free Speech
- Suppression of Freedom to Organize and Gather
- Threats
- Torture (in custody)
- Unlawful Arrest/Detention

Media

- Blog reports
- News/Media Reports
- Twitter Reports

Medical

- Disease/Sickness
- Health and Well-being Information/Programs
- Health Emergency/Crisis
- Health other
- Medical Supplies/Aid/Care Provided
- Vaccinations
- Women's Health and Reproduction (Includes pregnancy and childbirth)

Natural Disasters-Aftermath

- Death
- Displacement
- Houses/Property Damage
- Injuries
- Missing Person
- Trapped in Home
- Unstable Structure
- Animal Problems

Natural Disaster--Type

- Earthquake
- Fire
- Flood
- Hurricane
- Landslide
- Snow
- Tornado
- Tsunami

Political Organization

- Boycott
- Protests
- Strikes
- Military Defections

Prices

- Prices of Food and Commodities

Public Goods Provision and Infrastructure (NON-Emergency)

- Electricity
- Food
- Property/Houses
- Roads
- Sewer System and Sewage
- Trash/Garbage
- Water

Vacant Property

- Squatters
- Vacant Homes
- Vacant Lots

War

- Civil War
- Conflict with Foreign Country
- Intergroup Violence (Sectarian, ethnic, religious, etc. conflict WITHIN country)

Cities as Categories

- Cities (Code for Site Title!)

APPENDIX II: SURVEY RESEARCH QUESTIONS

EXPERIENCE

1. Before you used a Crowdmap, how would you describe your level of experience using this (or similar) crowdsourcing technology?

- No experience
- Some experience (1-2 times)
- Moderate experience (3-5 times)
- Considerable experience (6-10 times)
- Great deal of experience (more than 10 times)

2. How would you describe your level of experience as a contributor to digital maps?

- No experience
- Some experience (1-2 times)
- Moderate experience (3-5 times)
- Considerable experience (6-10 times)
- Great deal of experience (more than 10 times)

3. How would you describe your level of experience in hosting digital maps?

- No experience
- Some experience (1-2 times)
- Moderate experience (3-5 times)
- Considerable experience (6-10 times)
- Great deal of experience (more than 10 times)

4. What would you say was the most important reason why you engaged with Crowdmap:

1. I wanted to understand more about Crowdmap.
2. I wanted to create a map with a specific purpose and/or event.
3. I wanted to learn more about Ushahidi.
4. I wanted to demonstrate Crowdmap and/or participatory mapping to others.
5. Other - please specify: _____

5. If you setup a Crowdmap but did not actually use or create a map, please tell us why.

1. I was curious and did not intend to use a map.
2. I wanted to create a map but found it to be too technically challenging.
3. I wanted to create a map but realized it would take too much time.
4. I had security concerns about the collected data.
5. I wanted to create a map but Crowdmap was not the right tool.
6. I did not understand the processes for aggregating data from other sources (SMS, Twitter, email, etc.)
7. Other - please specify: _____

6. If you concluded that Crowdmap was not the right tool, what was the most important reason you reached this conclusion?

- It could not be customize to my needs.
 - It was too complex for my needs
 - It was too simplistic for my needs
 - Other - please explain briefly below:
-
-

7. If you did build a Crowdmap platform, which of these statements best characterizes your experience?

- I was able to generate the required public awareness of my map to make it as effective as I had hoped.
- I was able to generate the required public awareness of my map to make it almost as effective as I had hoped.
- I was not able to generate the required public awareness of my map to make it as effective as I had hoped.

8. To the degree I was successful in generating public awareness of my Crowdsourcing map, I did this mostly by

- Engagement with traditional news media (newspapers, radio, television).
- Engagement with community organizations, civic groups, and professional organizations.
- Engagement by social media, SMS, texting, etc.

9. If answered in the affirmative (options 'a' and 'b') in question six, I also generated public awareness of the Crowdsourcing map by:

- Engagement with traditional news media (newspapers, radio, television).
- Engagement with community organizations, civic groups, and professional organizations.
- Engagement by social media, SMS, texting, etc.

FUTURE USE

Please indicate all considerations that apply OR Please indicate the most important consideration

In the future, I would consider using Crowdmap if:

- It was easier to use.
- I identified a need.
- I had more time.
- I felt more confident about the security of contributors.
- It had better instructions.
- If technical support was offered.
- I would not try to use Crowdmap again.

When getting started with Crowdmap, I found it:

- Made sense and was easy to use.
- Was slightly confusing, but I figured it out.
- It just didn't make sense and I never got it to work.
- Hard to find the training, documentation.

DEMOGRAPHICS

Will you please tell us a bit about yourself?

Gender

- Male
- Female

Birth Year _____

Occupation _____

Education

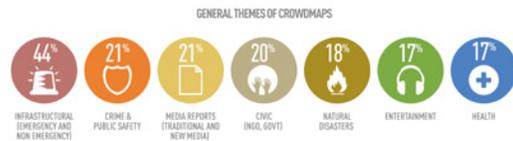
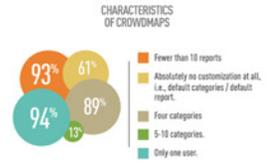
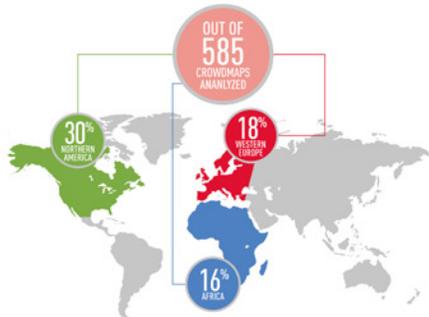
- Completed secondary education
- Some college
- College degree
- Some post-graduate education
- Post-graduate degree

In what city, country or region were you living when you used Crowdmap? _____

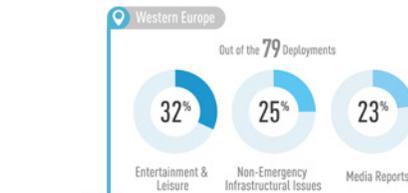
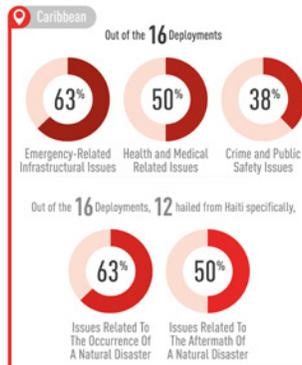
Was this also the location of the event or process for which you used or considered using Crowdmap?

- Yes
- No

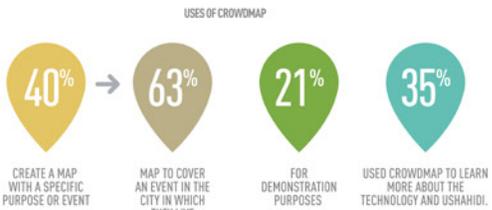
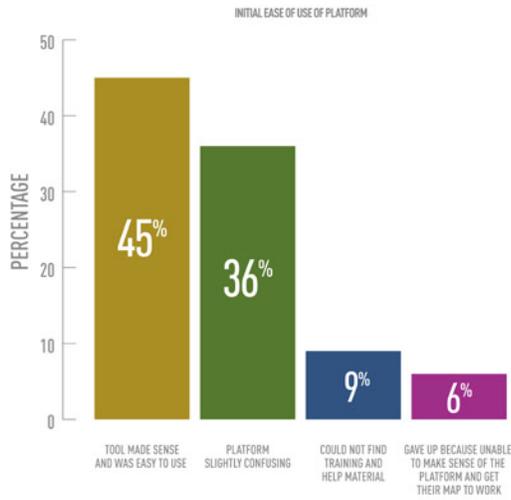
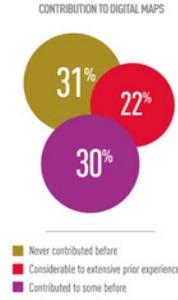
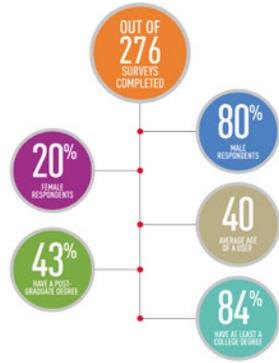
CROWDMAPS



REGIONAL COMPARISON BY POPULAR THEMES

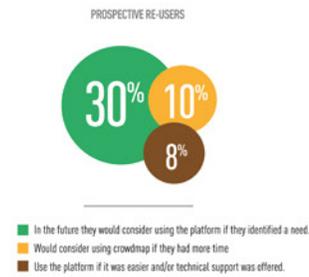
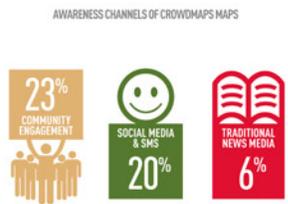
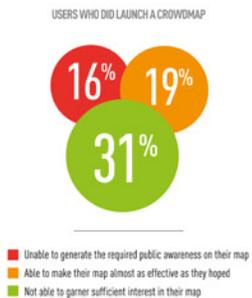
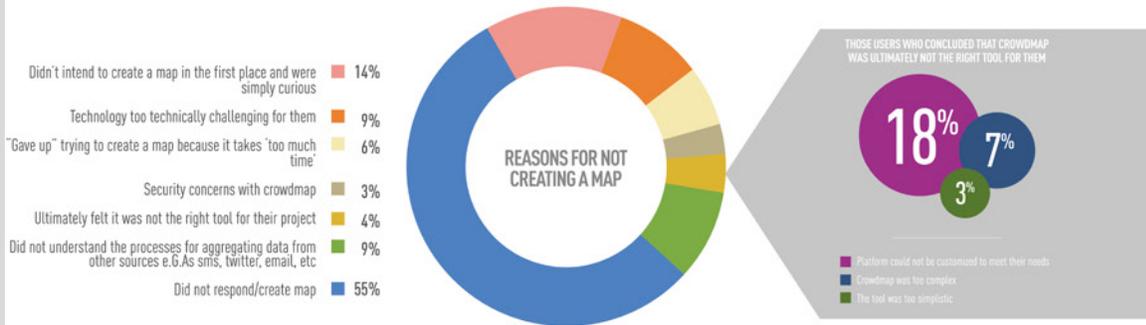


CROWDMAP USERS



IN OTHER WORDS, MORE THAN HALF OF CROWDMAPS WERE CREATED FOR NON-DIRECT USE.

MAJORITY OF CROWDMAP USERS WHO SET UP AN ACCOUNT DO NOT ACTUALLY END UP CREATING A MAP.



ABOUT THE INTERNEWS CENTER FOR INNOVATION & LEARNING

The Internews Center for Innovation & Learning supports, captures, and shares innovative approaches to communication through a creative program of research and development worldwide. Founded in 2011, the Center seeks to strike a balance between local expertise and needs and global learning in order to develop a comprehensive approach to understanding and catalyzing information exchange.

In Internews' 30-year history of promoting independent media in more than 75 countries around the world, the last five years have arguably seen the most changes in the global media and journalism environment. Across all Internews programs, adoption of cutting-edge technology is integral to advancing the work of the journalists, bloggers, citizen reporters, scholars and others who provide a vital interpretive role for their communities. The Internews Center for Innovation & Learning deepens and enhances our capacity to link existing expertise to research that helps define, understand and monitor the critical elements of changing information ecosystems and to pilot projects that apply and test the data, platforms and digital tools to meet information needs of specific communities. This is far from a solo endeavor. A network of partners, ranging from technologists to academics to activists is critical to creating and sustaining a dynamic and iterative collaborative space for innovation.

Internews Washington, DC Office
1640 Rhode Island Ave. NW Suite 700
Washington, DC 20036 USA
+ 1 202 833 5740

Internews Administrative Headquarters
PO Box 4448
Arcata, CA 95518 USA
+1 707 826 2030

www.internews.org
E-mail: info@internews.org
Twitter: @internews
facebook.com/internews

Internews is an international non-profit organization whose mission is to empower local media worldwide to give people the news and information they need, the ability to connect and the means to make their voices heard.

Internews provides communities the resources to produce local news and information with integrity and independence. With global expertise and reach, Internews trains both media professionals and citizen journalists, introduces innovative media solutions, increases coverage of vital issues and helps establish policies needed for open access to information.

Internews programs create platforms for dialogue and enable informed debate, which bring about social and economic progress.

Internews' commitment to research and evaluation creates effective and sustainable programs, even in the most challenging environments.

Formed in 1982, Internews is a 501(c)(3) organization headquartered in California. Internews has worked in more than 75 countries, and currently has offices in Africa, Asia, Europe, the Middle East, Latin America and North America.