Imagery to the Crowd, MapGive, and the CyberGIS: Open Source Innovation in the Geographic and Humanitarian Domains

By

Joshua S. Campbell

Submitted to the graduate degree program in Geography and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chairperson Jerome E. Dobson

Stephen Egbert

Xingong Li

Barney Warf

Edward A. Martinko

Date Defended: May 14, 2015

The Dissertation Committee for Joshua S. Campbell

certifies that this is the approved version of the following dissertation:

Imagery to the Crowd, MapGive, and the CyberGIS: Open Source Innovation in the Geographic and Humanitarian Domains

Chairperson Jerome E. Dobson

Date approved: May 14, 2015

Abstract

The MapGive initiative is a State Department project designed to increase the amount of free and open geographic data in areas either experiencing, or at risk of, a humanitarian emergency. To accomplish this, MapGive seeks to link the cognitive surplus and good will of volunteer mappers who freely contribute their time and effort to map areas at risk, with the purchasing power of the United States Government (USG), who can act as a catalyzing force by making updated high resolution commercial satellite imagery available for volunteer mapping. Leveraging the CyberGIS, a geographic computing infrastructure built from open source software, MapGive publishes updated satellite imagery as web services that can be quickly and easily accessed via the internet, allowing volunteer mappers to trace the imagery to extract visible features like roads and buildings without having to process the imagery themselves. The resulting baseline geographic data, critical to addressing humanitarian data gaps, is stored in the OpenStreetMap (OSM) database, a free, editable geographic database for the world under a license that ensures the data will remain open in perpetuity, ensuring equal access to all. MapGive is built upon a legal, policy, and technological framework developed during the Imagery to the Crowd phase of the project. Philosophically, these projects are grounded in the open source software movement and the application of commons-based peer production models to geography data. These concepts are reviewed, as is a reconception of Geographic Information Systems (GIS) called GIS 2.0.

Acknowledgements

First I have to thank my wife Hilary, for her love and support through this arduous journey. Without her, I could not have completed it. Our entire relationship has been in the shadow of this dissertation, and I look forward to our next chapter together.

This work would not have been possible without the dedication and support of a number of people. I would like to thank my advisor Dr. Jerome Dobson and the committee for their time and support for the duration of this process. I learned a tremendous amount from each of the committee members and their influence will continue to shape my professional life. Beyond being an advisor, I would like to thank Dr. Dobson for sharing his broad vision and passion for geography. His belief in the power of geographic technology, and vision for its role in science and society gave me the courage to dream audaciously. Next, I'd like to thank Dr. Lee Schwartz, The Geographer of the United States, for the opportunity to work at the U.S. Department of State. Dr. Schwartz provided the vision for the role of GIS in humanitarian emergencies, the platform for action in the form of the HIU, and the "top cover" to pursue these ideas. It was an honor to serve my country, and represent the diplomatic corps of the State Department.

My colleagues at the HIU were critical elements to the success of these projects. Patrick Dufour brought incredible technical talent and creativity, without him, none of these would have become a reality. Additionally, Christine Fellenz, LTC Karen Puschus, LTC Lauren Kulinski, Erin Sawyer, and Lisa Gregory all deserve credit for their efforts. From the Office of Innovative Engagement, Hilary Brandt, Janine Byas, Oliver Mertz, Nathan Kleekamp, Kate Henderson, and Deborah Eisenberg, were amazing collaborators and their efforts transformed Imagery to the

Crowd into MapGive. Honorable mention goes to Graham Lampa for connecting the pieces together.

Getting anything done in the federal government is a challenge. It requires taking risks and having people to back you up. To that end, there are a group of special individuals who were the team that pushed Imagery to the Crowd into a reality: Benson Wilder, Katie Baucom, Shadrock Roberts, and Chad Blevins, aka, the Original Master Cluster. Much love.

Finally, there would be no reason for Imagery to the Crowd and MapGive if it weren't for the Crisis Mappers community. The collective humanity, passion, intelligence, and creativity of this group provided the fertile ground for these ideas to take root. John Crowley deserves a special commendation; representing multiple stakeholders and communities that formed the critical nucleus around the Camp Roberts experiments. Additional thanks are due to the Humanitarian OpenStreetMap Team: Kate Chapman, Mikel Maron, and Schuyler Erle. And ultimately, it is the volunteers that make all of this work. Without the crowd, there is no crowdsourcing. This project is useless without the consistent, dedicated efforts of the unpaid volunteers that contribute their labor in service of helping humanity, proving time and again, their efforts make a difference.

Table of Contents

Introduction

Geography is a critical component in disaster response, disaster risk reduction, complex emergency response, and sustainable development. The ability to visualize changing events over space and time is a proven method for providing situational awareness and coordinating activity before, during, and after a disaster. The digital tools of geographic data collection, analysis, and visualization (Geographic Information Systems (GIS), satellite remote sensing, Global Positioning System (GPS), and cartographic visualization) are recognized as the required tools, but expensive proprietary solutions and the lack of trained staff still hinder widespread adoption in the developing world where both disaster risk and occurrence are highest (World Bank 2014). And while there is an abundance of evidence to confirm this, reports entitled "Successful Response Starts With A Map", "Geographic Information for Sustainable Development in Africa" (National Research Council 2007, 2002), and the Report of the Panel on United Nation Peace Operation (Brahimi 2000), indicate there are often gaps in geographic data that is vital for preparation and response. Specifically, it is the lack of digital geographic data that is structured, pre-processed, and formatted for a GIS that hinders the effectiveness of GIS in humanitarian applications (Currion 2006; Wood 2000).

However, this situation is changing. Advancements in internet connectivity, the proliferation of user friendly geographic tools, distributed volunteer communities, and open source software have begun to improve access to the power of GIS and geographic data collection. Empowered with these tools, new organizations are fundamentally changing the methods by which geographic data can be created, analyzed, visualized, and shared. This emerging collection of tools, methods, and organizational structures, termed GIS 2.0 herein, are the foundation for the development of a new methodology to catalyze the production of geographic data that can be used across the response and development spectrum. The MapGive initiative, deployed by the U.S. Department of State, leverages the notions inherent to GIS 2.0 in an effort to engage a globally distributed volunteer mapping network for the creation of free and open geographic data. The amount of data, and speed of production, from this volunteer mapping group is fundamentally changing the dynamics and economics of geographic analysis is disaster response and development.

MapGive is a unique development at the State Department, an organization traditionally more comfortable with narrative text and qualitative description than geographic data and visualization, but its utility has been recognized in two distinct manners. First, MapGive was selected by the State Department as a flagship initiative for the 2014 Open Government Plan (U.S. Department of State 2014), and second, it was referenced in the 2015 Quadrennial Diplomacy and Development Review (QDDR) in the Harnessing Knowledge, Data, and Technology section (U.S. Department of State 2015c, 53). However, MapGive did not emerge spontaneously; it was born from the vision for the role of Geographic Information Systems (GIS) in complex emergencies laid out by William Wood, former State Department Geographer.

Published in Geopolitics, Wood (2000) leverages experiences responding to the Kosovo conflict to describe the potential of GIS for complex emergency response planning and coordination. In a quote that presaged the need for MapGive, Wood (2000, p. 29) discussing the need for data in humanitarian interactive maps states "Unfortunately, such maps are often taken for granted and/or assumed to already exist, when more often they do not…" Recognizing there was a systematic gap in knowledge about complex emergencies, Wood and others at the State

Department created the Humanitarian Information Unit (HIU), the office I worked in for five years. The origin and mission of the HIU were used to understand the problem space of geographic data gaps in complex emergencies, and the HIU itself provided the organizational platform to build the geographic tools and processes required to address it.

The MapGive Initiative

This dissertation is about my five year effort to further embed the tools of geographic data, analysis, and visualization into the operational capabilities of the U.S. Department of State, specifically within the Humanitarian Information Unit (HIU), an interagency unit focused on the coordination and analysis of complex humanitarian emergencies. The crowning achievement of this work is the creation of the MapGive, Imagery to the Crowd, and CyberGIS initiatives. In short, these three initiatives represent an intentional strategy to combine the Web 2.0 concepts of network-enabled collaboration and peer-production of knowledge, with a high-performance geographic computing infrastructure built solely from open source software, to build a repeatable, sustainable process to address the problem of geographic data gaps in complex emergencies and humanitarian disasters.

The culmination of this work, the MapGive initiative, is a State Department branded project designed to increase the amount of free and open geographic data in areas either experiencing, or at risk of, a humanitarian emergency (U.S. Department of State 2015b). To accomplish this, MapGive seeks to link the “cognitive surplus” (Shirky 2011) and good will of volunteer mappers who freely contribute their time and effort to map areas at risk, with the purchasing power of the United States Government (USG), who can act as a catalyzing force by making updated high resolution commercial satellite imagery available for volunteer mapping. Leveraging geographic computing technology, MapGive publishes updated satellite imagery as standard compliant web services, essentially a standardized method of communicating map data across the network. These services can be quickly and easily accessed via the internet, allowing volunteer mappers to trace the imagery to extract visible features like roads and buildings without having to process the imagery themselves. The resulting baseline geographic data, critical to addressing humanitarian data gaps, is stored in the OpenStreetMap (OSM) database, a free, editable geographic database for the world being built by volunteers, largely from scratch (OpenStreetMap Wiki 2015e). OSM is often referred to as the Wikipedia of maps, and it uses a similar data license that ensures the data will remain open in perpetuity, ensuring equal access to all.

MapGive, however, does not accomplish all the mapping on its own, rather it participates as one element of a larger ecosystem built around the Humanitarian OpenStreetMap Team (HOT). A U.S. based non-profit organization, HOT is the community of volunteer mappers that utilize crowdsourcing and the OpenStreetMap database for the mission of "applying the principles of open source and open data sharing to humanitarian response and economic development" (OpenStreetMap Wiki 2015b; Humanitarian OpenStreetMap Team 2015b). The goal of MapGive is to help catalyze volunteer efforts through HOT by providing updated imagery for projects with a vetted humanitarian need, and to increase the number of volunteers by providing outreach, education, and training materials.

MapGive's goal of recruiting more volunteer mappers leverages two unique features of OSM mapping. First, since the satellite imagery services are delivered via the internet, there is no requirement that a volunteer be in the place being mapped. So long as they have an internet connection, a volunteer can contribute efforts to mapping tasks that could be located anywhere in the world. Second, OSM supports two data editing applications, one a stand-alone software install intended for advanced mappers or those who frequently work offline, and the second is intended to be more user friendly and runs in a standard web browser (Firebaugh 2013). This means that a HOT volunteer does not have to install and configure software to participate, and with a bit of training can begin mapping without needing a background in GIS.

This process of mapping from imagery, known as “remote mapping”, allows volunteers from around the world to collaborate simultaneously, producing map data at incredible rates and density. HOT also provides access to their microtasking platform, the OSM Tasking Manager, an application specifically designed to coordinate mapping volunteers by breaking down large mapping tasks into smaller, individual areas to map (Humanitarian OpenStreetMap Team 2015a). Mapping projects that utilize MapGive-supplied imagery services have produced data used by responders and development practioners in a range of situations and locations; the results of which are often stunning.

At the time of this writing there is a global effort underway to map the area affected by the Nepal Earthquake, with mapping contributions from volunteers from around the world. Statistics from May 3, 2015 indicate that 3,988 mappers have contributed to the effort, with 2,672 starting their OSM accounts specifically to help the response (Anderson 2015). Figure 1 below shows the temporal pattern and volume of volunteer mapping. Due to the large amount of imagery openly donated by commercial satellite imagery providers, MapGive supplied imagery was only needed in two areas. Figure 2 shows the map extents of the edits made from the MapGive-supplied imagery service.

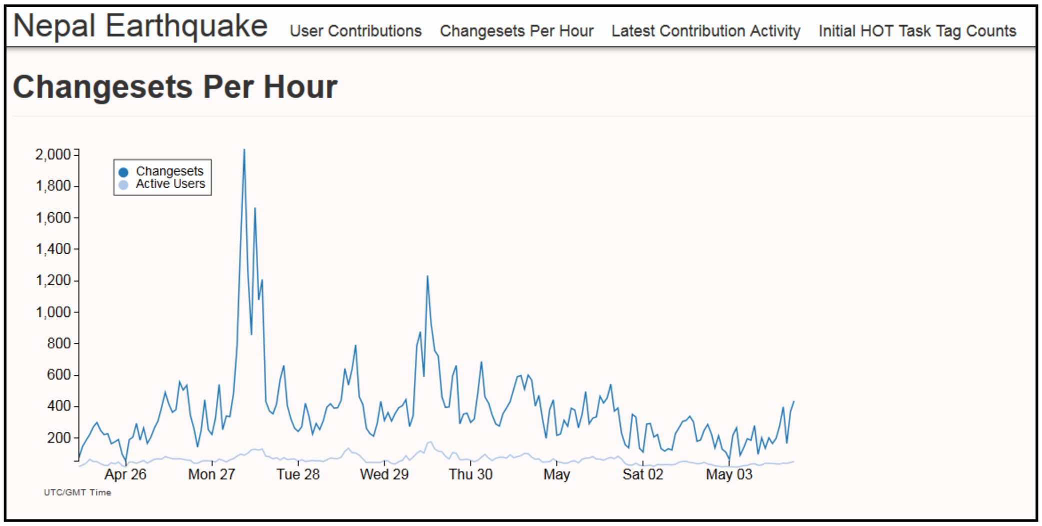

Figure 1 – Chart showing the number of active users and OSM changesets created during the HOT Nepal Earthquake response (Anderson 2015).

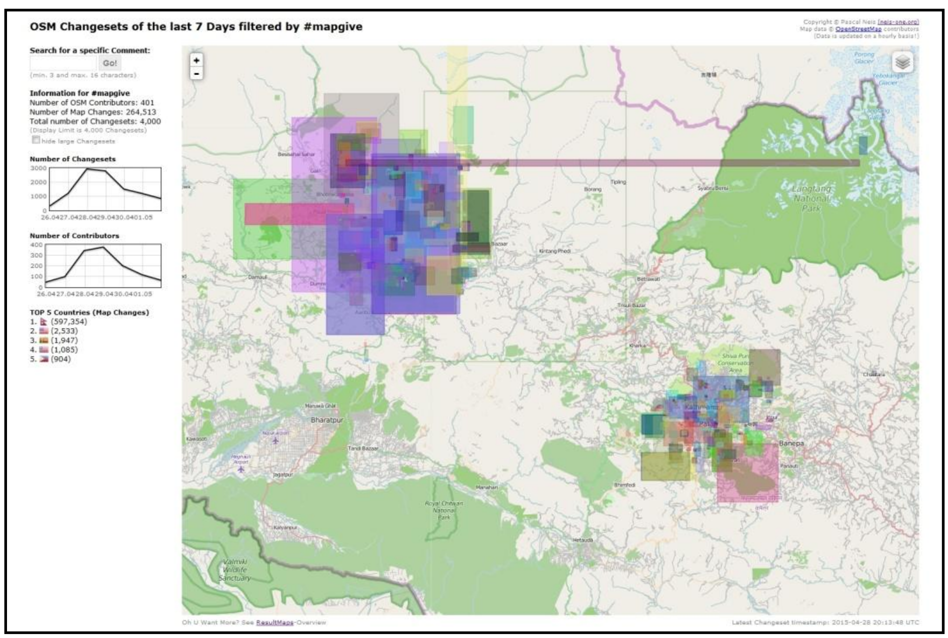

Figure 2 – Bounding boxes showing the extents of MapGive related map edits in support of the Nepal Earthquake between April 26 - May 2, 2015 (Neis 2015)

The MapGive initiative is actually composed of three distinct, yet inter-related components. At the foundation is the HIU CyberGIS, the project name for the geographic computing infrastructure built specifically for the computing needs related to the analysis and coordination of complex emergencies. I designed and built the CyberGIS using solely free and open source software, it is fully accredited for network security to run on both the internal State Department computer network and in a “cloud” deployment in Amazon Web Services. Second, the Imagery to the Crowd project was a multi-year effort to establish the legal, policy, and technical framework for sharing USG-purchased commercial satellite imagery with volunteer mappers. And third, MapGive is the educational and outreach components designed to exponentially scale the number of volunteer mappers participating in humanitarian mapping projects.

MapGive also links in with other elements of the State Department mission and was intentionally designed to connect the foreign policy objectives of supporting humanitarian crises and decreasing conflict with the public diplomacy goal of outreach to foreign populations. Since 2008 the State Department has spent significant effort building and maintaining a social media presence, with estimates in 2013 that the Department had approximately 30 million fans and followers across all its nearly 3,000 platforms and accounts (Brandt and Campbell 2013). This huge following has tremendous potential, and MapGive was an attempt to transform “social media followers into social media do-ers” by providing them the means to learn why mapping is important, how to map in OpenStreetMap, and to connect them with vetted mapping tasks known to have a humanitarian impact.

MapGive and Imagery to the Crowd are a data focused project built upon the technological foundation of the CyberGIS. However, the CyberGIS is designed to do much more, including the publishing of analytical interactive maps and spatial data catalog functions for sharing geographic data that can be used by the State Department to communicate and coordinate efforts internally and with external partners. The CyberGIS project borrows its name and design paradigms from the concept of “Geospatial Cyberinfrastructure” (Yang et al., 2010) and was built to support the needs of complex emergency analysis. Utilizing a modular architecture, the CyberGIS is built from solely free and open source software. This choice of open source was based on much more than simply cost, and was intended to be a means to collaboratively develop capabilities that could be freely shared with other government agencies, humanitarian organizations, and where useful, back to the open source software code repositories. Through the CyberGIS project, the HIU invested significant resources to add functionality to the open source mapping components as well as to openly publish software improvements made internally (Humanitarian Information Unit 2015b).

At their core, MapGive, Imagery to the Crowd, and the CyberGIS relied on new developments in geographic technology that have fundamentally changed the accessibility and usability of geographic data and analysis. Advances in internet connectivity and software development have led to online communities that collaborate in new ways, leading to a new set of conditions in which to deploy geographic tools. Through the lens of the Open Source Paradigm Shift I review how network-enabled collaboration, the peer production of knowledge, and software commoditization have created new opportunities for organizations to leverage the power of geography. These factors form the underlying strategy for the development of MapGive, Imagery to the Crowd, and the CyberGIS.

This paper is divided into six sections. In the first section, The Open Source Paradigm Shift, ideas on network collaboration and peer production of knowledge and their relationship to open source software and the economics of knowledge production are explored. The second section, GIS 2.0, reviews how technology advances and peer production systems have changed the application of geography. The third section discusses humanitarian data gaps, reviewing the requirements of Currion (2006) and Wood (2000) for incorporating GIS applications into complex emergencies and how those requirements helped found the Humanitarian Information Unit (HIU) at the Department of State. This section also covers the emergence of the Crisis Mappers and HOT. It was HOT's role in the 2010 Haiti earthquake response that changed the art of the possible in disaster response, and provided much of the inspiration for MapGive.

The fourth section reviews how the Imagery to the Crowd initiative was created, with a particular focus on the technology, legal, and policy challenges faced in developing an imagery sharing framework. The section includes results from five Imagery to the Crowd projects, each in a different geographic location and for a different purpose. The fifth section discusses the development of MapGive, the collaborative team, some of the successes of the project, and the educational and social media resources generated from it. This section concludes with a review of the outreach efforts to date and the role of mapathons in increasing the number of volunteer mappers. The sixth section discusses future research directions for these projects and reviews some of the issues addressed in the critical GIS literature.

The Open Source Paradigm Shift

Since 2005, the world of geographic technology has undergone tremendous growth, leading to the proliferation of online mapping services, an almost ubiquitous ability to access maps and navigation, and a renaissance of geography (or at least maps) in the public's mind. In the ten years since the launch of Google Maps, and the start of this revolution, there has been a convergence of several technological and social factors that have fundamentally changed the way geographic data is created, analyzed, visualized, and shared. The roots of this revolution lie outside of geography, found primarily in the world of telecommunications, information technology, and the new model for knowledge creation pioneered with the open source software movement. Understanding the dynamics behind this open source revolution is necessary for building the next generation of geographic applications.

The defining characteristics of open source software are ensconced in two related documents called the "Free Software Definition" and the "Open Source Definition" (Free Software Foundation 2015; Perens 1999; The Open Source Definition 2015). These definitions establish the principles by which source code will be created and distributed, a model that seeks to protect a user's freedom to run, copy, distribute, study, change, and improve the software. As Weber (2000, 2) summarizes, "the essence of open source software is 'free' – that is – open, public, and non-proprietary." To protect these freedoms, source code is released under specifically designed software licenses, which are also referred to as "copyleft" licenses. There are many open source licenses available, and parsing the differences is beyond the scope of this work, but the first and most famous of these in the GNU General Public License (Free Software Foundation 2007).

Software created using open source processes has proven to be extremely valuable, and is used to power the majority of the infrastructure of the internet, including domain name service (DNS), the Linux operating system, the Apache web server, various database, and multiple programming languages. In describing these projects Weber (2000, 35) states:

> Open source projects have demonstrated that a large and complex system of code can be built, maintained, developed, and extended in a non-proprietary setting where many developers work in a highly parallel, relatively unstructured way and without direct monetary compensation. The resulting piece of knowledge -- a technical artifact called software -- is remarkably powerful, efficient, and robust.

Beyond the source code and licenses, the true power of open source lies in the way knowledge is produced, and the social changes it precipitates when applied outside of software.

Coined by O'Reilly (2005), the Open Source Paradigm Shift is an attempt to explain how the mechanisms used by early software developers to build and share open source software became a social revolution that has extended to nearly every field of human endeavor. By casting his arguments in Kuhn's notions of a "paradigm shift", O'Reilly (2005, 480) argues that the concepts of open source extend significantly farther than source code and software licenses, that open source is the "natural language" of networked communities. In expanding this view, O'Reilly (2005, 480) is moving open source from a technological to a social phenomenon that can apply to any mode of knowledge production, and defining open source as a "field of scientific and economic inquiry, one with many historical precedents, and part of a broader social and economic story."

The basis for O'Reilly's claim of paradigm shift rests on a collective set of inter-related technological and social processes that emerged from building open source software, a process that involves a distributed, volunteer community of software developers working collaboratively to build a new generation of complex software by using the Internet as a means of communication. In the Open Source Paradigm Shift, O'Reilly's focus is primarily on the technological implications that emerged from widespread use of open source processes (software commoditization, network collaboration, and software customizability), and their impact on business models. The social implications of applying open source practices is limited in his review; however, other scholars also observed this relationship, further exploring the implications of network enabled mass collaboration and applying it to the production of knowledge and cultural goods. In expanding the view beyond software, Benkler (2002, 7) states:

> Rather than trying to explain what is special about software or hackers, I generalize from the phenomenon of free software to suggest characteristics that make large-scale collaborations in many information production fields sustainable and productive in the digitally networked environment without reliance either on markets or managerial hierarchy.

Benkler labeled these characteristics as the "commons-based peer production" model, taking the "commons-based" element from the open source practice of using licenses to ensure the freedom of source code and "peer production" to from the collaborative nature of the production. A comprehensive definition from Benkler and Nissenbaum (2006, 394) defines this model as a:

> "a socio-economic system of production that is emerging in the digitally networked environment. Facilitated by the technical infrastructure of the Internet, the hallmark of this socio-technical system is collaboration among large groups of individuals, sometimes in the order of tens or even hundreds of thousands, who cooperate effectively to provide information, knowledge or cultural goods without relying on either market pricing or managerial hierarchies to coordinate their common enterprise."

Benkler's recognition of the impact of network-enabled collaboration and definition as commons-based peer production models is a more robust interpretation of the mechanics of O'Reilly's open source paradigm shift. In a way these terms (OSPS and commons-based peer production) can be used interchangeably, with the distinction that O'Reilly was focused on the technological and business model implications of the mass collaboration, and Benkler on describing how it could apply in both technical and social domains. In either case, it is the connective power of the Internet to bring together distributed individuals into communities that precipitated these changes.

Paradigm Shifts

Originally published in 1962, Thomas Kuhn's The Structure of Scientific Revolutions introduced the term "paradigm shifts" into our cultural lexicon, and has since become an influential work on the dynamics of ideas in science (Kuhn 1996, xi). In Structure, Kuhn provides a theory to explain how new ideas in science come to replace old ones, a sequence by which scientific ideas are accepted (becoming a paradigm), are reinforced through experimentation designed to prove the paradigm, and how gaps, or anomalies, in the paradigm are eventually identified. The accumulation of anomalies causes the community to question the viability of the original paradigm, precipitating a period of "crisis". During the crisis period, experimentation continues until a new theory explains the anomalies missed in the first and the community "shifts" to accept the new paradigm, the period of "revolution".

While the domain of Structures is clearly the physical sciences, its model for the nature of paradigms and how they emerge, persist, and collapse can be applied to other domains. The rigor of scientific observation and crisp conclusions of mathematically derived formula often do not transcend into the social or cultural domains, but the nature of the battle of ideas does. Paradigms exist in any domain where theory is used to explain behavior, and in this case, the paradigms in conflict are the economic models for producing goods and knowledge. A review of open source software history, beginning with the release of the Linux operating system in 1991, through the

2000s and the emergence of Web 2.0, to today's hyper connected world of mobile and smart devices, reveals several of Kuhn's principles at work and serves as a timeline for defining the milestones of the open source paradigm shift.

The two competing paradigms in this battle are traditional firm-based production model and the commons based peer production. As Benkler (2002) explores in "Coase's Penguin", the traditional models of production relied on either market pricing or firm-based hierarchical project management as the forces that organized the production of goods. But with the introduction of digital networks, a third model based on the notion of commons-based peer production emerged. In this model, ad hoc networks of distributed actors participate cooperatively in the production of knowledge without the hierarchical control of the firm model and without the pricing signals provided by the market (Saveri, Rheingold, and Vian 2005). Peer production models run contrary to the existing firm and market models for two reasons. First, activity is self-organized in a bottom up approach, and second, because the motivations of participants in these systems is not driven by market pricing and monetary reward, a characteristic that confounded many economists.

These dueling models of knowledge production also form the central argument of Eric Raymond's work (2001) on this era of open source software development entitled "The Cathedral and the Bazaar". Raymond contended that the firm model resembled the top-down processes of building a cathedral: guarded, structured, and integrated, whereas open source was more like the bottom-up approach of the bazaar: chaotic, rapid, and iterative. Besides highlighting that the open source method was dramatically different than the standard, commercial approach to software development, this book described how the inclusion of collaborative volunteers, connected over the internet, produced better software because the code was open and could be reviewed.

Continuing with the paradigm shift metaphor, it was trends in the software market during the 1980s towards increasing proprietary control over source code that highlighted an "anomaly" within the firm model of software production. This anomaly violated the open sharing of source code that was an accepted process in the culture of computer scientists, a practice that was technologically required during the mainframe computing period before it became a philosophical principle. Upon the introduction of wide area networking, individual software developers who held the view that source code should remain open were suddenly able to work together, collaboratively, to create software in a non-proprietary model. The fact that developers participated in these collective projects without any monetary reward presented an additional anomaly in the firm based production model for software.

When the Linux operating system was released in 1991, it exposed the failure of the firm-based production model to explain why open source practices were capable of producing complex and powerful software. Kuhn describes this as the "crisis" period, when previous theory no longer explains a phenomenon. The resolution of the crisis phase occurs when a new paradigm emerges, and by Weber (2000) and Benkler (2002), the commons based peer production model provided a viable alternative. With the emergence of successful business models built from the principles of peer production and other "technologies of cooperation" (Saveri, Rheingold, and Vian 2005) around the turn of the millennium, the revolution phase of the paradigm shift had begun.

In a more structured analysis, Baldwin and von Hippel (2010), in an article entitled, "Modeling a Paradigm Shift", compare three types of innovation models, including the producer (analogous to the firm model), single user, and open collaborative (analogous to the peer production model). Their analysis sub-divides each innovation model into a set of “technological properties” composed of four “costs”: design, communication, production, and transaction. These costs can be estimate for any given product, and depending upon the relative breakdown of the costs, the utility of a given innovation model can be assessed. Essentially this framework provides a mechanism for estimating the impact that a given process change or new technology may have on the optimum innovation model, providing the ability to model how the growth of internet connectivity and decreasing compute costs affect software production.

During much of the 20th century, the dominant mode of innovation is the producer’s model. In this model, producers are motivated to develop innovations in which they can sell back to customers, in the form of goods and services. Having the potential for multiple customers allowed firms to spread the cost of development over a large base. This model reinforced the system of Intellectual Property, in which the firm has legal protection over its ideas, thereby protecting its ability to sell or license its products without fear of copying. In this case the high cost of designing new products and coordinating teams to work on the product (communication cost) tipped the economic conditions such that the producer model was most viable.

However, with the growth of internet communication and ever decreasing compute costs, both design and communication costs have decreased, resulting in an increase in the utility of single user and open collaboration models. As noted in both Benkler and O’Reilly’s analyses, these technological features precipitated the emergence of peer production models in software production. In Baldwin and von Hippel’s framework these technological advances and concomitant reduction in costs have allowed the single user and open collaboration models to become more viable across the economic spectrum, allowing them to compete with the producers model in many sectors of the economy. Quoting them (p1399):

> "The shift to the single user and open collaboration innovation models is driven by new technologies – specifically, the transition to increasingly digitized and modularized design and production practices, coupled with the availability of very low cost, Internet based communication."

Baldwin and von Hippel's conclude that the technology driven shift from producer to user and open collaboration innovation models is profound enough to be considered a paradigm shift, echoing O'Reilly's analysis.

Kuhn states that it can take an extended period of time from the initial crisis, to the emergence of a new theory, through the end of the revolution and a new paradigm is dominant. It took approximately 15 years until the release of Linux translated into successful Web 2.0 business models built on open source practices of network enabled collaboration and commons-based peer production. Paradigm shifts take so much time because they ultimately result in a changed world view, meaning previous scientific inquiry on the subject has to be reconsidered. Quoting Kuhn (p7):

> "That is why a new theory…is seldom or never just an increment to what is already known. Its assimilation requires the reconstruction of prior theory and the re-evaluation of prior fact, an intrinsically revolutionary process that is seldom completed by a single man and never overnight."

Additionally, Kuhn defines an additional characteristic to determine if a paradigm shift has occurred (p11):

> "Their achievement was sufficiently unprecedented to attract an enduring group of adherents away from competing modes of scientific activity. Simultaneous, it was

> sufficiently open-ended to leave all sorts of problems for the redefined group of practitioners to resolve."

Determining if an "enduring group of adherents" exists requires looking at open source in a few different ways, including as software, as knowledge production, and as business model. It is clear that open source software, and the legions of developers who contribute their efforts on GitHub, BitBucket, and other source code management platforms, meets this criteria. While statistics are difficult to obtain, GitHub reports there are 22.4 million individual code repositories on their platform (GitHub 2015), though it is unclear how many of these are private. In terms of peer production of knowledge, the most significant application of open source methodologies beyond software is Wikipedia, a web-based, free-content encyclopedia project. Wikipedia has consistently grown since its introduction in 2001, with over 4.8M English language articles, 25M registered users, and almost 800M edits (Wikipedia 2015b). Figure 3 shows the trajectory of growth in English language articles, a curve that is still increasing at an impressive rate, and that clearly indicates a large number of people are still contributing to and using the system.

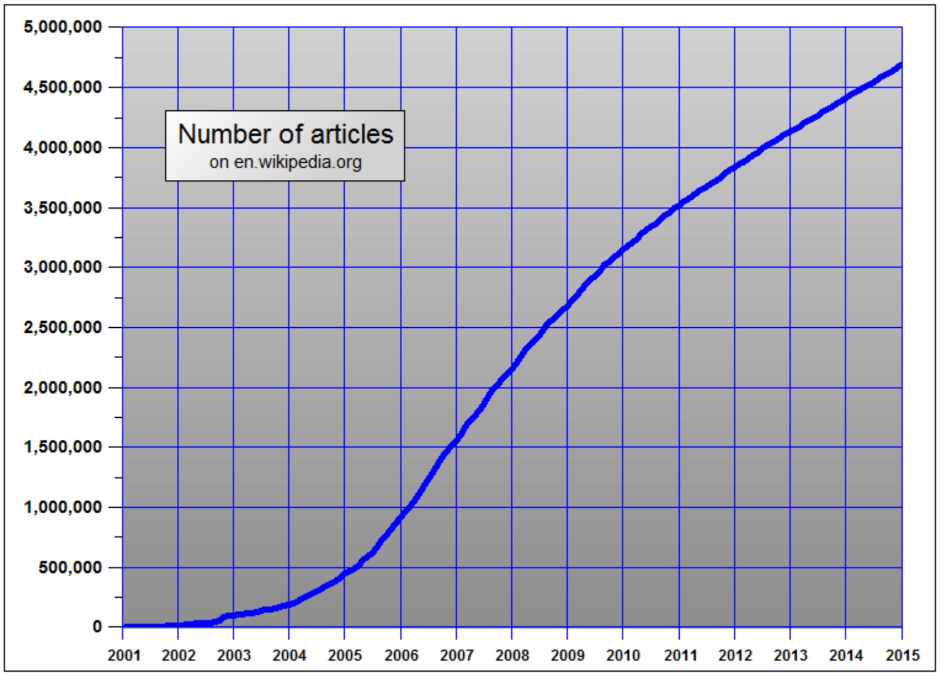

Figure 3 – Cumulative total of English language Wikipedia articles ("EnwikipediaArt" by HenkvD - Own work. Licensed under CC BY-SA 3.0 via Wikimedia Commons - http://goo.gl/637pCK)

As the business models that emerged out of Web 2.0 extend into their second decade, we can evaluate their performance. Names like Google, Amazon, Facebook, and eBay are all essentially for-profit business models that harness the output of network-enabled communities. In their own way, each one of these companies relies on their ability to aggregate the actions of individual users on their platform, then analyze that content and present it back to users in support of their efforts to search the web, find products and reviews, or connect with people. Their impact is so large that we now cannot imagine a past without them, a corollary to Kuhn's point about paradigm shifts causing total re-evaluation of previous fact. There is a cadre of venture capitalists that are essentially the new adherents to the peer production business model, and as the endless array of social media and file sharing applications demonstrate, they are actively applying this new paradigm to every domain possible, determining the limits of success and failure points of the paradigm. Many of the new "collaborative economy" companies like Uber, AirBnB, and Instacart represent the next wave of this approach, taking network-enabled collaboration models from the knowledge domain, adding geographic location, and applying it to problem of optimizing the delivery and usage of physical goods. These businesses, themselves built from open source software and internet collaboration, have thrived in the marketplace, displacing many of the old guard companies and our notions of how companies should operate.

Taken collectively, the academic research on innovation models, the success of for-profit business models and the continued growth in open source software developers and Wikipedia articles demonstrates that "open source" is more than a software phenomenon. Open Source as a collaborative practice has fundamentally changed the way products and knowledge are designed and built, attracting an enduring group of adherents over the 25 years since its introduction to society, and hence qualifying for the designation of a paradigm shift.

Implications for Geography

Geography as a science, geographic information as digital data, and geographic functions implemented as software all exist within the broader domain of knowledge that has been impacted by the Open Source Paradigm Shift. The introduction of network enabled collaboration, open source software development, and peer production models have all had an effect on geography, increasing the availability of geographic software and introducing a new model for geographic data production.

With the first releases of the MapServer and GeoServer web mapping servers around 2000, open source began to make its mark on geography and web mapping. Over the next several years the pace of development and deployment increased substantially, driving a process of commoditization across the "stack" of geographic computing elements (spatial database, web mapping servers, and client rendering libraries). Combine that software with commoditized "cloud"-based computing models, and the cost to deploy an enterprise-grade geographic computing infrastructure is measured in pennies per hour. The process of commoditization was increased by with the establishment of Open Geospatial Consortium (OGC) standards for encoding geographic features in a spatial database (Simple Features Specification) and the communication of geographic data over the network (Web Mapping Services). The introduction of OGC standards brought interoperability into geographic computing and had the knock-on effect of increasing the adoption of open source components as modular systems built from a hybrid of open and proprietary software applications could be assembled together.

As predicted by disruptive innovation theory and the law of conservation of attractive profits (Christensen and Raynor 2003), a process of commoditization does not destroy value, instead, commoditization drives a concurrent process of de-commoditization at an adjacent tier or sub-system of the technology stack. Value migrates out of the commoditized tier into the adjacent de-commoditized tier, in this way value isn't destroyed, just transferred. The classic example of this process is the standardization of IBM personal computer hardware interfaces in the 1980s, leading to the rapid commoditization of once profitable computer hardware. As the cost of computer hardware came down dramatically, value shifted to the subsystems of the computer, mainly the central processing unit and the operating system. Out of this process, Microsoft and Intel reaped tremendous benefits.

By 2005 and the release of Google Maps and Google Earth, the commoditization of the geographic software tier began to drive value into the adjacent data tier. Data refers here to both the vector data used for cartographic basemaps, navigation routing functions, and point-of-interest (POI) searches, and the raster imagery used as a visual basemap. Companies like TeleAtlas and Navteq built billion dollar businesses by collecting, cleaning, and reselling road and POI data, while DigitalGlobe and GeoEye did the same for satellite imagery. The major commercial map providers, Google, Bing, Yahoo and MapQuest, were all licensing data from these companies.

O'Reilly (2005) also cites disruptive innovation theory and the "law of conservation of attractive profits" as the best way to explain the impacts of commoditization caused by from the introduction of open source software. In most Web 2.0 business models, a peer production model is utilized to offset the cost of building a proprietary database. For example, Amazon uses an open source software stack to run their datacenters. Value in this case flows from the commoditized tier, computing, into the de-commoditized adjacent data tier, information about books and other products. Amazon built tremendous value from their ability to present detailed book reviews and ratings for nearly every book, and to tell a user that individuals who purchased the book they are viewing, also bought some other book. To build the database for support this, Amazon essentially used a peer production system to crowdsource reviews and ratings.

O'Reilly (2005, 473) specifically references geographic data and mapping applications as the "counter-example" to overcoming de-commoditization, as no geographic applications had yet to leverage the power of peer production models to address the data cost issue. O'Reilly (2007) repeats this notion when discussing Web 2.0 business models. This means that by 2007 one of the brightest minds in Silicon Valley had repeatedly published the idea of a huge market advantage to those who could apply the principles of the Open Source Paradigm Shift to geographic data. This somewhat shocking reality highlights a strange feature of geographic technology, it tends to run 5-10 years behind mainstream IT. I have not been able to determine why this is true, but O'Reilly highlights it, and it is commonly accepted in the geographic software industry. However, this could be viewed as a positive, as there are advantages that can be gleaned by determining trends and processes that succeed in traditional IT, and deploying the geographic corollary to it.

Setting aside the data issue for a moment, it is important to remember that the introduction of Google Maps in 2005 did launch geographic technology into the Web 2.0 era and kick off a decade of tremendous growth. The "slippy map" interface of Google Maps and "spinny globe" of Google Earth fundamentally changed the user experience of online mapping, and set a new model for mapping interfaces. However, the true revolution emerged when developers figured out how to integrate their own data with a Google basemap, creating the first map "mash-up". Underlying the mashup concept was a technology built using web services where a user could publish their own data (photos, GPS tracks, favorite coffee shops, etc) in a simple data format that contained a structured geographic coordinate (Turner 2006). The Google Maps application programming interface (API) could then consume that data format, extract the location, and place a hypertext icon as an overlay on the basemap depicting the location and linking to the source data. The mashup, combined with GPS-equipped mobile devices, exposed the power of mapping to everyone in a simple way, and is what drove the popularity of mapping with average citizens.

There have been many names given to the phenomenon of non-geographers using online mapping tools to create custom maps, including neogeography (Turner 2006), volunteered geographic information (Goodchild 2007), popular geographics (Dobson 2012) and the wikification of GIS (Sui 2008). My point in mentioning these is not to review the large body of existing literature on this subject, as Haklay et al. (2008) "Web Mapping 2.0: The Neogeography of the GeoWeb" does a tremendous job, but to highlight the role that mashups had on allowing distributed communities to collaborate over the internet using geographic data. By simplifying the process of adding geographic data to user's content, applications like Flickr opened the door for individual photo contributions to be aggregated and shared from a central location an under a license that defined the level of sharing. In this way the map mashup became the tool that enabled a geography-focused peer production model for cultural knowledge. This process was what O'Reilly had indicated was missing from geographic applications.

However, the popularity of the mashup had a downside. It instilled a false dichotomy in users minds that mapping was about putting simple dots, that they controlled, on top of complex, pre-rendered basemaps, that they did not control. As Haklay et. al. (2008) point out, the entire process of neogeography relied upon the rich geographic database provided by the large corporate mapping providers for geocoding and routing. The seeming ubiquity of the Google basemap, and the fact that it was basically free to use, means that for most people and most applications, there was a tremendous amount of highly curated base data instantly available. However, this instantly raises a problem for humanitarian applications. What do you do when Google does not have data?

This brings the discussion back to geographic data and O'Reilly's point that geography had not applied peer production methodologies for data creation. However, that wasn't entirely true, as the OpenStreetMap (OSM) project a " free, editable map of the whole world that is being built by volunteers largely from scratch and released with an open-content license" had begun in 2004 (OpenStreetMap Wiki 2015e). OSM is discussed in additional detail in a later section, but for now it is important to recognize that OSM is the geographic data extension of the Open Source Paradigm Shift. Through its peer production model, volunteer mappers contribute geographic data collected from GPS tracks or by tracing imagery and upload it to a centralized server. As with other peer production systems, volunteers participate for the same varied and non-monetary reasons. . OSM has experienced tremendous growth since its introduction is now a locus of tremendous value. And as I'll show with MapGive, when there is a true humanitarian need for geographic data, OSM contributors can create data at a rate that has never been seen before.

This section has attempted to review the complex dynamics and inter-relationship that have emerged as distributed communities became linked by the internet. The peer production collaboration model and new digital tools that emerged organically from this increased level of communication have precipitated a paradigm shift in how we conceive of the production of knowledge across all domains and economic endeavors. Geography is no exception to this, and with an understanding of these dynamics, it is possible to craft the next generation of geographic applications that leverage their capabilities.

GIS 2.0

When originally proposed, this dissertation intended to focus on a new model of Geographic Information Systems (GIS). Termed GIS 2.0, it was anticipated this reformulation of GIS would share key similarities to the reformulation of the Internet that became known as Web 2.0, web centric, service oriented architectures, peer production systems harnessing collective intelligence, data focused, and user experience focus. The origins of my thinking began in the mid-2000s, when the rate of innovation in the geographic space was exploding. With the Google Earth and Maps launch in 2005, and the first Apple iPhone in 2007, suddenly geography and maps were accessible everywhere and usable by everyone. By 2010, when I was finally able to begin work directly at the State Department, it was no mystery that geography was being dramatically affected by the internet, and there were more a number of new names to describe it. The concepts of WebGIS, the Wikification of GIS, Volunteered Geographic Information (VGI), Virtual Earth, Digital Earth, Neogeography, Web Mapping 2.0, Popular Geographics, and the GeoWeb began to occupy the semantic space used to describe the combination of Geography and Web 2.0 concepts I had envisioned.

Now in 2015, it is redundant to predict the emergence of a reformulated GIS, it already happened. The implications of the combination of Web 2.0 and geography have been discussed in the literature for many years, with VGI and GeoWeb appearing to be the lasting terms of art. And while the name at this point is almost irrelevant, I believe there are still enough differences between what I termed GIS 1.0 and GIS 2.0 to warrant discussion. Therefore the review presented here is not designed to recap every work along the way, but rather to highlight the concepts of GIS 2.0 that provided the technological structure I used to build the CyberGIS and MapGive.

Using an analogy from the Web 2.0 definition, GIS 2.0 "doesn't have a hard boundary, instead it is defined by a gravitational core"(O'Reilly 2007). Figure 4 is a graphic representing the 'memes' that compose the gravitational core of GIS 2.0. At the center is the set of requisite core competencies that I believe organizations which implement GIS 2.0 must have:

- Geographic analysis
- Widespread data collection
- Open Source
- Collective Intelligence
- Network Effects
- Mobile
- Grassroots

Surrounding the core is a set of technological themes, innovation models, and social contexts that are contributing elements to GIS 2.0. With few exceptions, each of the contributing elements represents some facet brought on by either the open source paradigm shift or the concept of the web as a platform. As O'Reilly (2007, 34) stated about Web 2.0, many of the technologies that underlie it were not new, but the combination "represented a fuller realization of the true potential of the web platform." This same logic applies to GIS 2.0, but in a different way. For GIS 2.0, most of the geographic theory and GIS functions are not new, but the combination of geographic principles with web technology designs and open source processes has lead to a "fuller realization of the potential of geographic analysis."

This notion of GIS 2.0 was built specifically with the analytical issues related to humanitarian emergencies and the HIU in mind. And while most of the themes and contexts could apply to any domain, the focus on global issues and populations at risk are key elements of the HIU's analytical portfolio. Similarly, the core competencies section represented the central tenets of what I believed the HIU needed organizationally to leverage the capabilities of GIS 2.0. The design of the CyberGIS had to take into account the type of applications and analyses it would be used for, as well as the organizational culture that was trying to use it. For four years I kept a copy of this graphic tacked above my monitors at the HIU, as it is the blueprint for what I was trying to build.

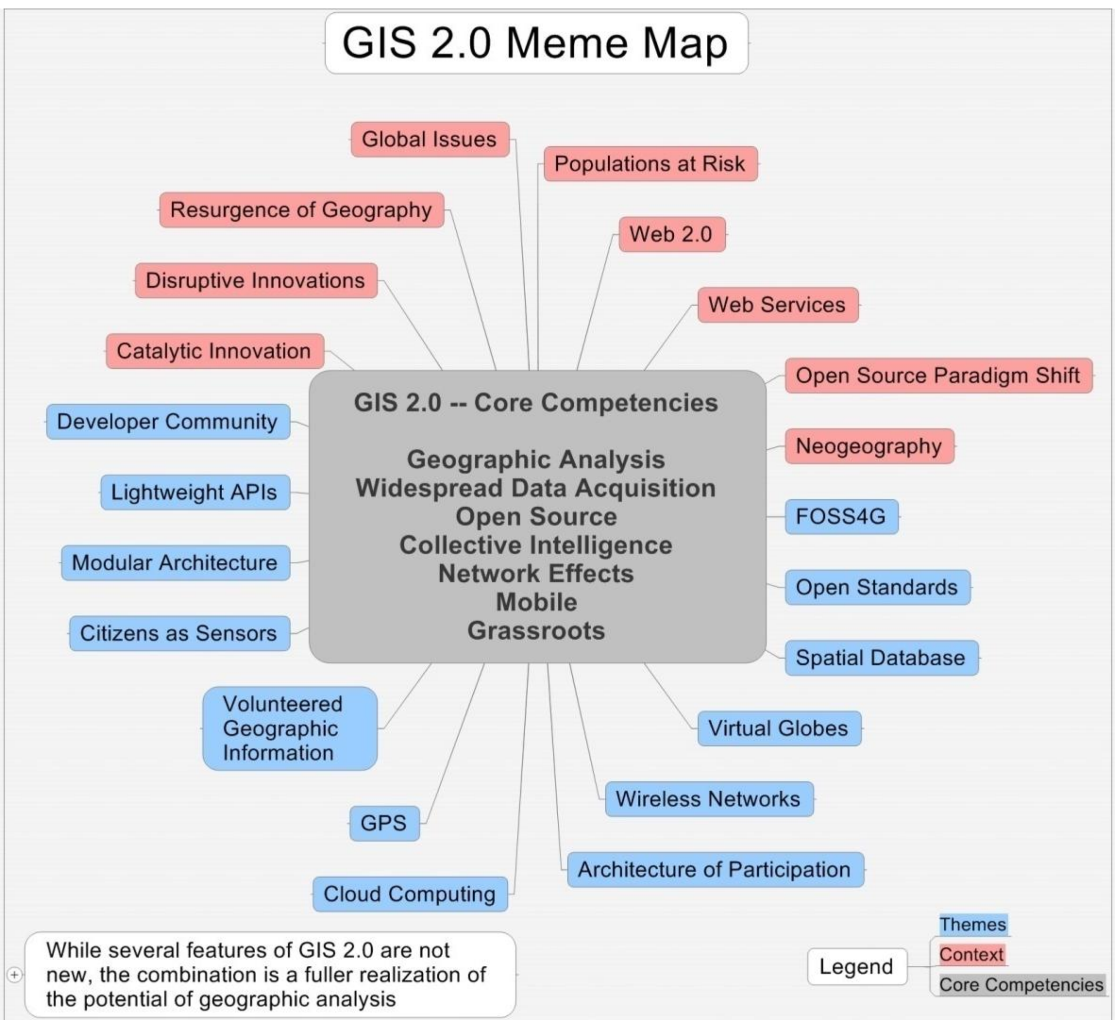

Figure 4: GIS 2.0 Meme Map composed of core competencies (center, grey), technological themes (bottom half, blue), and philosophical contexts (upper half, red).

Technology

The intersection of the web, open source, and geographic technology had a tremendous impact on the mechanisms for implementing geographic functions. At a conceptual level the impact has been a subtle and consistent diffusion of GIS functions from stand-alone, local systems into the distributed, network-enabled IT mainstream, what some have called Spatial IT (Ramsey 2014). This diffusion involved a technological shift from specialized analytical and data handling software, bespoke geographic formats, and local computing systems, to the integration of geographic data and functions into standard IT tools, formats, and workflows, with connectivity and reach extended from the desktop to be distributed on the web. Using a modified version of Clarke's (1995) definition of GIS, Table 1 lists the characteristics of how each GIS function was implemented in GIS 1.0 vs GIS 2.0.

GIS Roles (modified from Clarke 1995)	GIS 1.0	GIS 2.0
General Trends	High capital costs, required significant expertise, local computing systems, proprietary software, publishing model	Near-zero acquisition cost, increased usability, web service architecture, virtualized, cloud, mobile, open source software, participation model
Collection	Survey, digitize, Stand alone GPS, Authoritative data	GPS-enabled mobile devices, High resolution commercial satellite imagery, peer production, Natural language processing
Storage / Retrieval	Flat files, dbf files, local system	Spatial RDBMS & NoSQL, cloud computing, distributed
Integration	Manual	Semantic Ontology

Analysis	Stand alone software and algorithms, algorithms and data on local systems	Integrated into database functions (SQL), distributed data and web services (WPS) , MapReduce / Hadoop, GeoTrellis, GeoMesa
Visualization	Paper maps, on screen, primarily 2D	Web distributed, accessible by multiple devices in multiple formats, web browser interfaces (virtual globes, 'slippy' maps)
Dissemination	Hard copy maps, physical media	Digital maps and data, spatial data catalogs, catalog services for web standards (CSW), distributed web services (W*S)

Table 1: Comparison of GIS 1.0 and GIS 2.0 traits

Of all the technologies listed in Table 1, there are a few that should be highlighted as particularly diagnostic of the GIS 2.0 transition from desktop to web.

- Spatial Database

The development of the Simple Features standard by the Open Geospatial Consortium (OGC) and International Organization for Standardization (ISO) defined a standard encoding format for geographic objects, such that they could be stored as a native data type in a relational database and manipulated using structured query language (SQL) (Open Geospatial Consortium 2007). Additionally, standard GIS vector overlay operations are also included and accessible via SQL. From a philosophical level, the migration of standard GIS functions into the database and accessible via standard SQL processes marks a tremendous step towards the native incorporation of GIS into standard IT. PostGIS, the open source spatial extension for the PostgreSQL database fully implements this standard. As part of the CyberGIS project, the HIU invested in the completion of PostGIS version 2.0, which also incorporated raster analysis functionality (PostGIS Project n.d.).

- OGC Web Services

Collectively the Open Geospatial Consortium created a series of standards that defined interfaces for communicating geographic data across the web. These include standards for raster, vector, imagery, and geoprocessing functions. These standards provide a mechanism by which a client application, usually a web browser, can make request for geographic data using a defined structure, encode that request in a URL that gets passed over the network to a web mapping server. The server can then parse the URL request to determine the geographic layers, extent, and map projection. The role of standards is a powerful force in both Web 2.0 and GIS 2.0 in that they provide interoperability, allowing software created by different organizations to communicate across a common interface.

- Powerful web browsers

As the computing aspect of GIS is pushed to the web, standard desktop / thick client installations of software are decreasing and the web browser is becoming the client of choice. Advancements in browser technology has allowed for larger and more complex geographic applications to be consumed via the web. At the HIU we had to wait until the Department approved Google Chrome for use on OpenNet, the Department's intranet, before we could deploy any our applications. Before that point, only Microsoft's Internet Explorer was installed and it simply could not handle the processing load of the application.

- Growth in GPS and commercial high resolution satellite imagery

While not internet related, the transition to GIS 2.0 could have occurred without the development of these two technologies. Specifically it was the inclusion of GPS chips into mobile phones that enabled georeferenced mobile data collection and location-based services. The development of a commercial remote sensing capability that had 1-meter and less spatial resolution introduced a new mechanism for the creation of spatial data. I think it is important to highlight the role the United States government played in the development and commercialization of these two technologies. In two different policy decisions implemented by President Clinton, Selective Availability on the GPS signal was removed on May 1, 2000 significantly increasing its locational accuracy and precision, and restrictions limiting the development of high resolution commercial satellites were removed, leading the new generation of satellites orbiting now (Berger 1994).

As discussed in the previous chapter, the Open Source Paradigm Shift resulted in an inter-related set of technical and social processes. As such, GIS 2.0 was also envisioned to be more than just a technological implementation, that it would leverage the social features of peer production of knowledge. Building a true GIS 2.0 system at the HIU required the integration of two existing projects representing the technological and social element: Geospatial Cyberinfrastructure and OpenStreetMap.

Geospatial Cyberinfrastructure

The term “cyberinfrastructure” emerged in the late 1990s in the context of the critical infrastructure underlying cyberspace, and has over time been defined generically as the combination of “data resources, network protocols, computing platforms, and computational services that bring people, information, and computational tools together to perform science or other data-rich applications” (Yang et al. 2010). In a whitepaper entitled “Cyberinfrastructure Vision for the 21st Century” the National Science Foundation (2007) defined a new model of computing for scientific discovery built from the concepts of cyberinfrastructure (CI). The report documented four guiding principles that would guide NSF's investment in cyberinfrastructure: 1) high performance computing, 2) data, analysis, visualization, 3) virtual organizations for distributed communities, and 4) learning and workforce development. As expressed in these two definitions, cyberinfrastructure encompasses a wide array of computing, analytical, and organization components necessary to collect, store, analyze, model, visualize, and disseminate data.

The Internet plays a central role in cyberinfrastructure as it connects distributed database and analytical engines via web services into a seamless computing environment. Similarly, the notions of network-enabled collaboration are built into the framework of cyberinfrastructure in the form of tools for community development and collaboration. Cyberinfrastructure has been extended into the geographic domain by including specialized components for handling geospatial data, spatial analysis and modeling functions, and cartographic visualization (Yang et al. 2010; Wang 2010). Referred to as Geospatial Cyberinfrastructure or CyberGIS, these computational frameworks are often, if not exclusively, built from open source software, and are predicated on network-enabled collaboration for both people and compute functions. The term CyberGIS has also been used as a project name for an NSF-funded project that is researching the technological and social elements of spatial problem solving (Wang et al. 2013). Geospatial cyberinfrastructure is the closest technological instantiation and semantic description of the requisite technological elements I had envisioned for GIS 2.0, and as such, the system that I built at HIU was entitled the CyberGIS.

OpenStreetMap (OSM)

Started in 2004, the OpenStreetMap project is a “ free, editable map of the whole world that is being built by volunteers largely from scratch and released with an open-content license” (OpenStreetMap Wiki 2015e). In many ways OSM is the geographic equivalent to Wikipedia, where a user can create an account and begin to contribute information. In the “Technologies of Cooperation” nomenclature (Saveri, Rheingold, and Vian 2005), OSM would qualify as both a peer-production network and a knowledge collective, meaning it leverages crowdsourced efforts of individual mapping edits and aggregates them into a centralized database. Volunteer mappers contribute digital geographic data in vector format obtained from GPS devices or from tracing features (roads, rivers, buildings) from imagery. Attributes describing elements of the features (road name, business name, address) are also collected. OSM requires that for a feature to be included in the database it must be physically observable. For instance demographic data at an aggregated level would not be appropriate for inclusion.

Since its beginning in 2004, OSM has seen tremendous growth, having recently exceeded 2M registered users (Figure 5)(OpenStreetMap Wiki 2015c). And at just over 2.8 billion nodes, OSM has become an incredible resource for geographic data, often the de facto map in areas of the world where there is limited commercial motivation to produce geographic data. OSM has been the focus of a significant amount of research in geography, a trend I don’t see slowing down. Neis and Zielstra (2014) provide a deeply researched and comprehensive survey of the history, recent developments, and future research directions with OSM. The relevance of OSM in the context of GIS 2.0 is its open source founding principles, organizational structure, and potential application in the humanitarian domain.

OSM has proven, just as the open source software movement proved with the release of Linux and Apache, that a distributed, volunteer community utilizing a commons-based peer production model is capable of producing and maintaining a complex, sophisticated body of knowledge. Metrics for data quality, accuracy, and completeness in OSM obviously vary by location, but by 2009 OSM had grown enough in many parts of the United Kingdom and western Europe that the data was becoming commensurate in terms of quality and quantity of that produced by national mapping agencies (Haklay 2010). In Germany, where OSM is extremely popular, it compares quite favorably to a commercial dataset, missing only 9% of features, but as the OSM mapping includes a number of sidewalks, paths, and trails not included in the road dataset, it exceeds the total number of segments by 27%. Neis and Zielstra (2014) provide an extensive review of several data quality analyses.

Organizationally, the OSM Foundation is a non-profit foundation “supporting, but not controlling, the OSM project” (OpenStreetMap Foundation 2014). Neither OSM nor the OSM Foundation has any paid employees, which is amazing given the technological skills needed to maintain the network and server infrastructure of the system. Data stored in OSM is protected by data-specific “copyleft” license called the Open Database License (Open Data Commons n.d.), a variant of the type of license applied to open source code, which ensures the data will remain free and open.

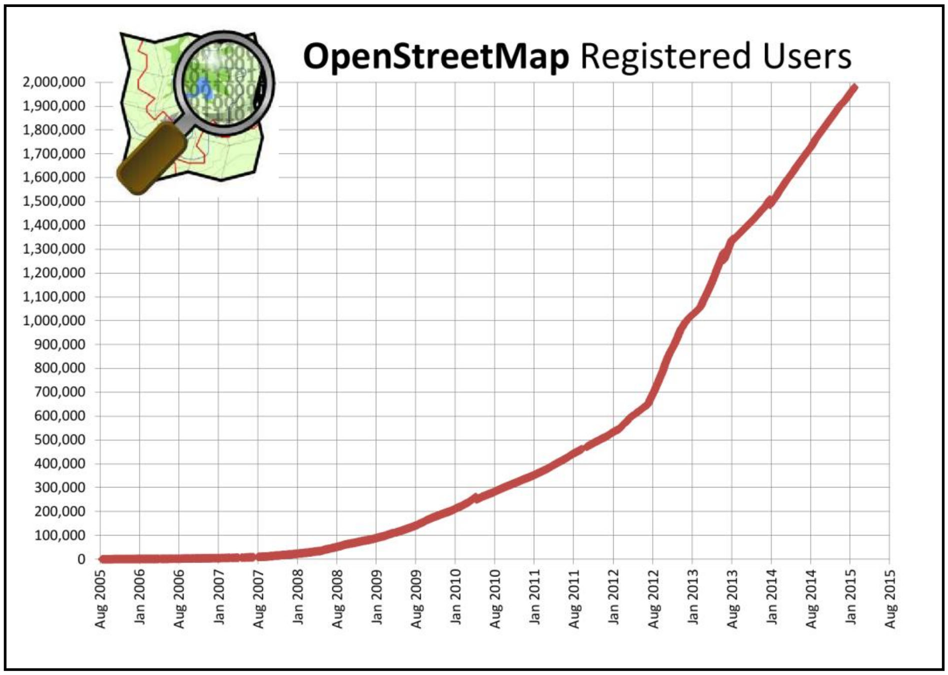

Figure 5 – Growth curve of number of registered users in OpenStreetMap

Imagery Services and Remote Mapping

In the first several years of OSM the majority of mapping was through on-the-ground surveys conducted by volunteers carrying GPS devices (walking, biking, and driving). While effective at generating quality data, using GPS track as the primary method of data collection required that the mapper physically be in the location they were mapping, a constraint that limited mapping to accessible and permissive areas. This changed with the widespread introduction of OSM mapping from satellite imagery.

On July 12, 2007, the internet company Yahoo became the first company to allow OSM users to use their imagery service for tracing and extraction of vectors into the OSM database (Maron 2007; Yahoo! Aerial Imagery n.d.). This was a significant development as it was the first instance of a commercial venture (Yahoo) that was able to navigate the legal constraints under which they licensed the imagery to allow OSM tracing. Licensing is a complex legal issue, and one the OSM community spends a significant amount of energy on (OpenStreetMap legal-talk Archives n.d.). As such, it is beyond the scope of this work to cover all issues, save those broad questions related to imagery licensing and its impact on Imagery to the Crowd and MapGive (discussed later). As it relates to OSM the tricky legal question is whether the company can allow user to trace geographic features from the imagery and then store those extracted vectors in the OSM database and associated open data license.

A few years later, on November 30, 2010 Microsoft opened access to their Bing imagery holdings for tracing by the OSM community (OpenStreetMap Wiki 2015a). This was a tremendous boon to OSM mapping as the Bing coverage was significantly larger than Yahoo's and at finer spatial resolution in many populated area. As reported by Neis and Zielstra (2014) the introduction of Bing imagery greatly increase the amount of building mapping in OSM, a feature type that is often useful in humanitarian applications. From personal experience I've determined that Bing maintains a high standard for the georectification accuracy of the images they use in the map service, and I have a reasonably high confidence that the data extracted from their service is reliable.

On April 9, 2014 DigitalGlobe announced they were partnering with MapBox to allow OSM tracing right to the imagery service they had already been providing to MapBox (Bullock 2014). The MapBox service acts as a good compliment to the Bing imagery as often the scenes at a given location are different, so a volunteer mapper can choose the better option. The MapBox service has one potential drawback in that users are limited in how far they can zoom in, which is occasionally less than what is needed for certain tasks. Given that DigitalGlobe recommends a mapping scale of 1:12,000 based on their ortho-rectification processing pipeline, this limit may be tied to that scale.

Personal experience processing DigitalGlobe imagery has also demonstrated that georectification accuracy is highly affected by the off-nadir angle of the collection. DigitalGlobe satellites have the ability to look in either direction by a fair margin, which dramatically shortens revisit time, but there is an accuracy trade-off in collecting data in that manner. The reason the imagery in the Bing service is more consistent, even though it is also Digital Globe imagery, is that they only accept images with a very low off-nadir angle.

In this chapter I have attempted to extend the Open Source Paradigm Shift into a vision for GIS 2.0, defining a set of technological themes and philosophical contexts that combine into a set of core competencies that an organization using GIS 2.0 needs to embrace. Specifically tailored for humanitarian tasks, this vision of GIS 2.0 is built upon a set of technologies that are moving traditional GIS into a network-enabled, collaborative model. To build GIS 2.0 will require the computing architectures of geospatial cyberinfrastructure and the network enabled data collection processes of OpenStreetMap volunteers combined with satellite imagery providers.

Humanitarian Data Gaps

The previous section attempted to establish a framework for understanding that a new geographic toolkit emerged from the combined effects of the open source paradigm shift and advances in geographic data collection from GPS and high resolution commercial satellite imagery. The combination of these forces represents a powerful new capability for the discipline, so the question obviously becomes, for what purpose?

It has been long recognized that data gaps in emergencies are a serious hindrance to effective action. Whether it be the "fog of war" in the military context or it's corollary the "fog of humanitarianism" (Weiss and Hoffman 2007), the inability to take collective action due to a lack of information is a fundamental and persistent problem. Geographic Information Systems (GIS) have been demonstrated to be an effective tool for coordinating, analyzing, and visualizing data for situational awareness and decision support in humanitarian emergencies (Verjee 2005, 2007; Cowan 2011), peace keeping operations (Brahimi 2000; Currion 2006), and complex emergencies (Wood 2000). But even with widespread recognition of the value of GIS, Currion (2006) points out, there are many impediments slowing broader GIS adoption, including the lack of integration of GIS processes into operational workflows and the inability of GIS staff to execute at the necessary tempo to demonstrate their value during a humanitarian emergency.

Wood (2000) also highlights the need to integrating GIS into operations before the disaster, and to prepare data and applications before the disaster. In further discussing the role of GIS for coordinated responses in complex emergencies, Wood (p29) makes several key assertions:

1. GIS-based ‘humanitarian planning map(s)’ are a critical component for integrating disparate sources of information.
2. To be effective, geospatial data and digital planning maps need to be built and integrated into the planning process prior to the intervention.
3. To be useful these planning maps must be clearly organized, standardized, and fully GIS-compatible.
4. Dynamic base maps must form the core of an information package that interveners bring with them to the response.

At the time that article was written, William Wood was the Chief Geographer at the State Department and working on humanitarian issues in Kosovo. This combination of needs formed some of the requirements that Wood and others used to propose and build the Humanitarian Information Unit (HIU) at the U.S. Department of State.

Humanitarian Information Unit

Established in 2003, the Humanitarian Information Unit (HIU) resides in the Office of the Geographer and Global Issues within the Bureau of Intelligence and Research at the Department of State and serves as an interagency center that collects, analyzes, and disseminates information that helps U.S. Government decision makers and partners anticipate and respond to humanitarian crises worldwide (U.S. Department of State 2013). To accomplish this mission, the HIU utilizes a geospatial perspective and Geographic Information Systems (GIS) tools as a fundamental component of its business processes. As a result, most HIU products are map-based infographics designed for senior policy makers at the State Department. These map products are usually formatted for a single sheet of paper and provide a snapshot view of a given situation, see Figure

6 (U.S. Department of State 2015a). While these single page products are useful for senior policy makers, HIU experience has proven that many users at the State Department and across the interagency would benefit from having access to HIU data and products in a dynamic and interactive environment.

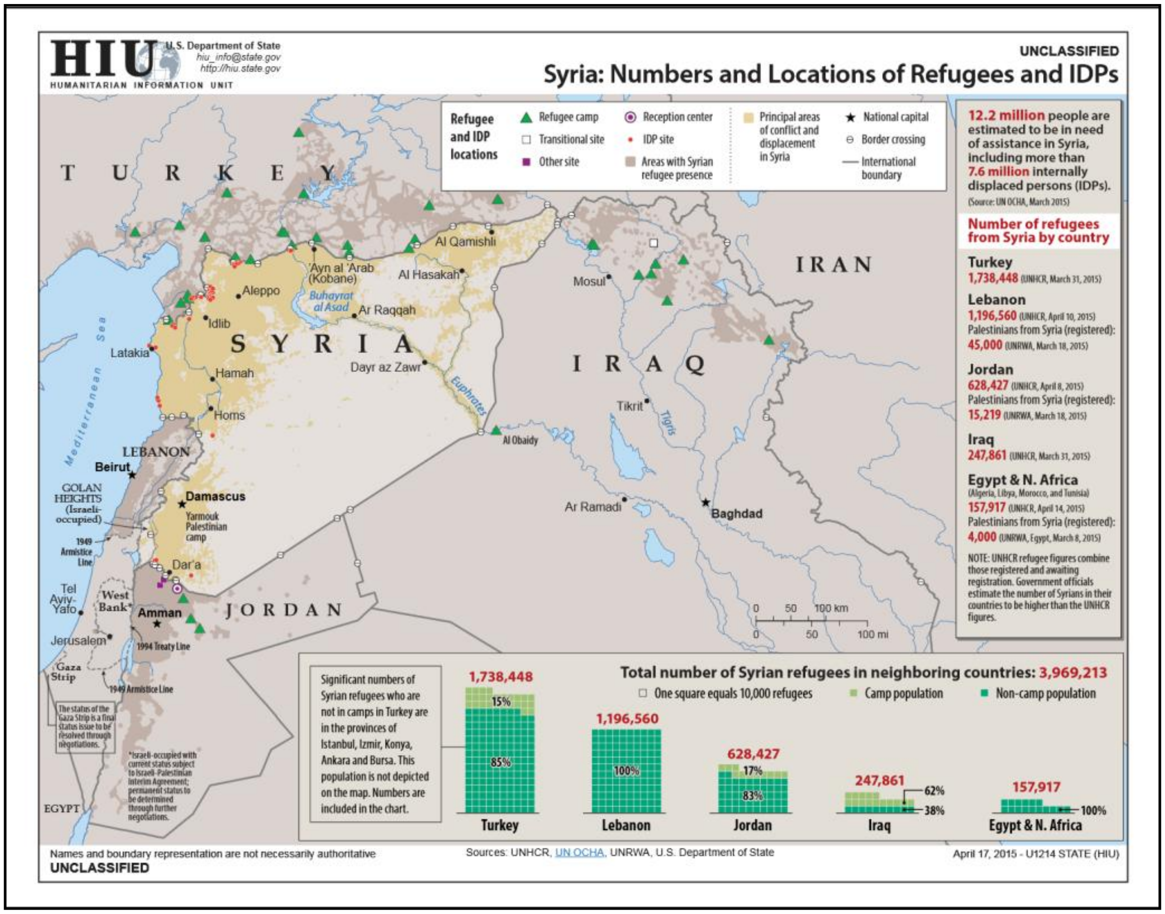

Figure 6 – Example of a typical HIU product

During my five years at the HIU, I worked to increase the efficiency of the existing product line, with the goal of extending HIU products into a web format that could be easily updated and provide users an interactive view. In many ways the HIU faces the same criticisms that Currion (2006) illustrated. GIS had been identified as useful tool, but maps were never explicitly integrated into State Department practices, so when the HIU was tasked for a product it could be on a very short timeline. Given there is always some lead time needed to obtain and process data for a GIS analysis, this often meant the HIU struggled to move at the high tempo of humanitarian emergencies, frustrating HIU analysts and policy customers alike. Ultimately the HIU had to get more agile, and similar to Currion and Wood's recommendations, the solution was to do as much as possible to prepare systems and data ahead of time.

This process started with a substantial revision to the HIU's internal data management structure and GIS workflow, restructuring it for consistency and reusability. Poor data management practices were a significant drag on efficiency as it was difficult to identify data at the start of a project, or to understand how a project was constructed when it needed to be updated. These changes had a demonstrably positive impact on the production workflow as it introduced consistency through all phases of the GIS workflow, from data acquisition though analysis, cartography, dissemination, and archiving.

The next step was to address the systems side of the equation, effectively how to graft a web-based publishing model on the existing HIU workflow. To do that we had to build a system, the choice was to pursue a new geographic computing infrastructure designed from the principles of GIS 2.0. This project came to be known as the HIU CyberGIS.

CyberGIS

The Humanitarian Information Unit (HIU) Geospatial Cyberinfrastructure (CyberGIS) is a web platform for transforming the HIU's map capabilities into interactive and exploratory mapping applications. This platform delivers geospatial data, analytical web services, and integrated mapping applications through a range of applications and formats. From a technical perspective, the platform is built from free and open source software, supports existing open web standards, utilizes a service-oriented architecture, and is deployed in a virtualized environment for datacenter and cloud deployment. Funding to build the CyberGIS was obtained through two proposals to an internal State Department competitive grant pool called The Innovation Fund. CyberGIS Phase I was awarded the largest amount of any grant given and resulted in a set of scalable, extensible, and reusable geospatial software tools that can be were accredited for use on OpenNet, the State Department intranet, and that can be freely shared within the Department, broader USG, and humanitarian community. CyberGIS Phase II followed on the success of Phase I and obtained additional funding for the GeoNode spatial data catalog.

As a web platform, the CyberGIS combines the visual power of mapping and Geographic Information Systems (GIS) with the rapid dissemination capabilities of the Internet. This combination, termed Geospatial Cyberinfrastructure or CyberGIS (Yang et al. 2010; Wang et al. 2013), was designed as a natural extension of the HIU's existing map product line, extending it into an interactive web environment. Beyond simply replicating HIU products on websites, the CyberGIS unlocks new ways of displaying and interacting with humanitarian information. The goal of the CyberGIS is to increase the HIU's capacity for data processing, spatial analysis, visualization across space and time, and communication with policy makers about the geographic dimensions of complex emergencies.

Underlying the HIU CyberGIS are four technological principles: Free and Open Source Software, Open Standards, Service-Oriented Architecture, and Cloud deployment. Designing the CyberGIS around these principles is a cost-effective and innovative solution that increases the efficiency of HIU processes, expands access to products, and leverages existing technology investments. At a higher level, the HIU CyberGIS supports a number of HIU and State Department strategic goals related to humanitarian analysis and information technology, and demonstrates Department of State (DoS) leadership in the Federal Open Data movement, the Free and Open Source Software movement, and the respective DoS and Federal Cloud Computing Strategies.

Defined generically, a CyberGIS is a web-based collection of geospatial functions, implemented through a set of technologies that enables users to obtain, analyze, and visualize geospatial information. The proposed HIU CyberGIS follows this pattern, and separates the implementation of these functions into three general technological categories:

1. Mapping server and spatial database
2. Browser-based mapping applications
3. Spatial data catalog

Web mapping applications are the primary way users interact with the CyberGIS. Delivering these applications requires development on both the server and client side. On the server side, the web mapping server and spatial database work together as a data store, map rendering engine, and analytical toolbox. The client side is a web browser application focused on user interaction and provides functions for map display, map rendering, map interaction, and other interface functions. Each of the software elements in the stack are open source projects that are bundled into a single deployment known as the OpenGeo Suite Enterprise Edition. Each of the various software components in the OpenGeo Suite are proven open source software projects with extensive developer communities. OpenGeo has built the Enterprise Suite to work as a single install package with a predictable update schedule, and keeps key developers from each project on staff. The OpenGeo Suite contains the components required for the web mapping server (PostgreSQL, PostGIS, GeoWebCache, and GeoServer), and the client-side web mapping applications (OpenLayers and GeoEXT libraries), see Figure 7. The mapping server and spatial database will be built as a VMWare virtual machines and Amazon Machine Instances (AMIs) with an Ubuntu Linux operating system, Tomcat servlet container, and the Apache web server software.

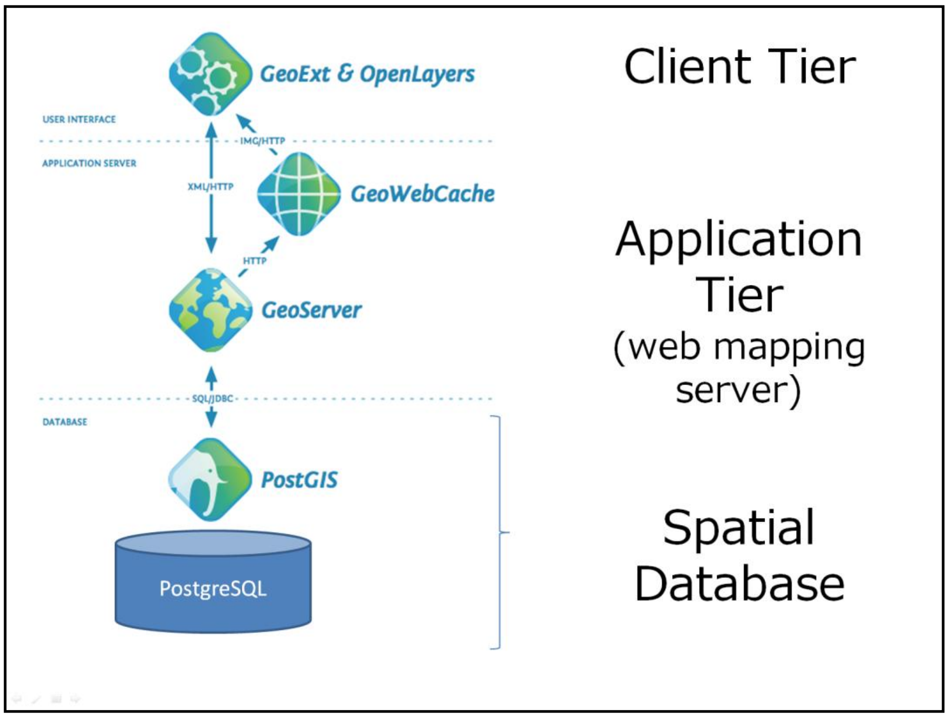

Figure 7 – Diagram displaying a simplified CyberGIS architecture and the open source components used in each tier

The third element of the CyberGIS, the GeoNode spatial data catalog, is a web application that allows a user to search, visualize, and download geospatial data. The HIU CyberGIS will make relevant data available through this application in a range of formats and web services. The spatial data catalog is built from the same suite of technology listed above, plus additional libraries that provide the catalog service and the Python-based web framework Django for the web interaction tier. All data loaded into the spatial data catalog will have International Organization for Standardization (ISO 19137) compliant metadata

Beyond the software itself, web services are the other key enabling technologies used in the HIU CyberGIS. Underlying the entire system is a service-oriented architecture (SOA) design pattern, with specific support for the Open Geospatial Consortium (OGC) collection of web service standards. Widely utilized in Web 2.0 mapping applications, SOA technology allows OGC web services to be remixed into various applications. Also known as “mash-ups”, these applications can integrate web services published from a range of providers. A single application can contain one or many different services, and in a CyberGIS application can contain both mapping data and analytical web services. These different types of services can work together to produce integrated “dashboard” applications.

Additionally, web services can be consumed by a range of software clients beyond web browsers. In this case, HIU web services can be consumed by professional-level desktop GIS software (both open source and proprietary), Google Earth, and mobile applications. Initially the system would “push” data to these clients, but with additional development it is possible to have clients contribute information back to the system.

The true innovation of this project is not in its technical specifications, open source software, or cloud deployment. Instead the “disruptive innovation” of the HIU CyberGIS is in leveraging the geographic dimension of data as an analytical framework and medium of communication. At its core, the State Department is organized geographically, but it is currently ill-equipped to leverage its own geographic information. Christian (2007) describes the limited use of GIS within the Department, despite widespread interest in the technology. DoS information is inherently place-based, but without the inclusion of maps with data, that information is translated solely via unstructured text written as cables or emails. By removing the geographic component from reporting, the information becomes very difficult to overlay with other data. Combining text reporting with structured geographic data means multiple streams of information can be visualized together and analyzed across time and space. The innovation of the HIU CyberGIS is beginning the process of DoS harnessing its geographic information for analysis and visualization.

The HIU CyberGIS reinforces a number of strategic goals ranging from the HIU itself, up through the Department and the broader USG. Below is a summary of strategic goals taken from the HIU mission statement and several other relevant strategic plans at the time when the grant proposal was written:

HIU Mission Statement:

> #1: Analyze and disseminate unclassified information critical to USG decision makers and partners in preparation for and response to humanitarian emergencies worldwide
>
> #2: Promote best practices for humanitarian information management

Office of the Geographer and Global Issues:

> Support INR mission as the Executive Agent for Outreach at the Intelligence Community
>
> Support development of Web 2.0 applications to more effectively link various humanitarian communities of practice

Bureau of Intelligence and Research (INR) Strategic Goals FY2012:

#3: Create and maintain an expert workforce

#4: Support the provision of humanitarian assistance

#5: Provide timely and focused all-source analysis related to the promotion of peace and security

DoS/USAID Joint Strategic Goals 2007-2012:

1: Achieving Peace and Security: Conflict Prevention, Mitigation, and Response

#5: Providing Humanitarian Assistance: Prevent and Mitigate Disasters

Department of State IT Strategic Plan 2011-2013 Goals:

#1: Digital Diplomacy – Collaboration, Information and Integration

#2: Cloud Computing

#3: IT Leadership

Federal Cloud Computing Strategy 2011

The National Intelligence Strategy 2009:

Intelligence Community Enterprise Objective #4: Improve Information Integration and Sharing

National Security Strategy 2010 Goals:

Security: Invest in the Capacity of Strong and Capable Partners

Prevent the Emergence of Conflict

Values: Promote Dignity by Meeting Basic Needs

CyberGIS Applications

The following section reviews three of the CyberGIS applications with the goal of displaying the potential of geographic visualization combined with structured data collection and thoughtful user interface design for understanding the spatial and temporal dimensions of a humanitarian emergency. Additionally, I hope these applications demonstrate that geospatial cyberinfrastructure can be used to address several of the requirements put forth by Wood (2000) and Currion (2006) and make the deployment of GIS-enabled applications more relevant to both policy makers and humanitarian responders. I'd also like to note that several other CyberGIS applications that were built during my time at HIU that were equally as innovative as what is discussed below, but that they are not reviewed here as they were built for internal purposes only.

Daily Humanitarian News Brief

The first application of the CyberGIS was built as a graphical interface for an existing project called the Daily Humanitarian News Brief (DHNB). One of the first changes I made at the HIU was to create a database reporting structure for the DHNB where HIU researchers could index important news stories of the day. An email is sent out every morning with a summary of these stories. The DHNB application combines a map interface with geocoded news events and allows a user to explore the last six weeks of news events by date, keyword, or location, Figure 8. Additionally, the entire DHNB archive is available on the HIU data page (Humanitarian Information Unit 2015a).

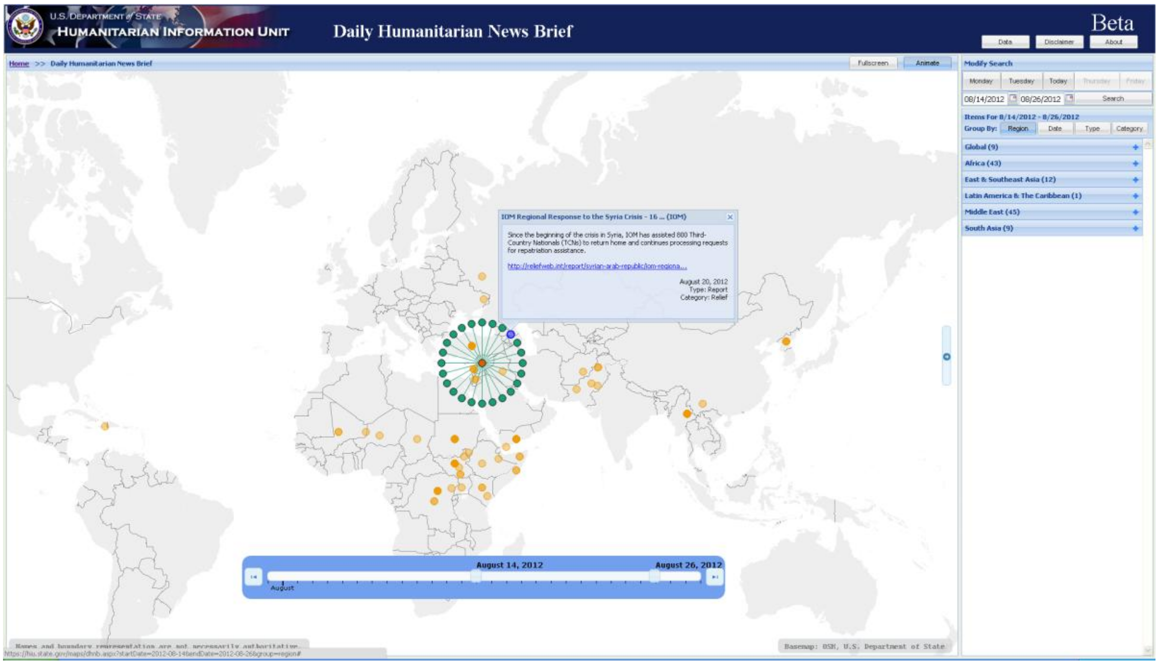

Figure 8 – Screenshot of the Daily Humanitarian News Brief web application

Horn of Africa (HoA) Viewer

The first integrated CyberGIS application was built as a visualization platform to explore the Horn of Africa famine of 2011-2012. In it users can interactively explore the geographic and temporal components of famine zones, vegetation health, livelihood zones, curated news reporting, and refugee data to understand the relationships between transnational food security issues and political, economic, demographic, and physiographic variables. The user interface was designed to be as sleek as possible, or as I said at the time “Google Maps simple”. There are none of the traditional GIS toolbars or map interface controls. There are two basemaps, the Blue Marble satellite layer and OpenStreetMap, and then a series of overlays that can be selected from the Map Layers dialogue box, Figure 9.

The most innovative part of the HoA Viewer interface is the Time Slider element at the bottom of the map view. Since the HoA Viewer is built on several time-series datasets, we needed an intuitive way to communicate to the user what time period they were viewing as well as the number of other time segments of that data are available. When a new layer is selected from the Map Layers box, several elements of the map automatically update. The Legend updates to display the relevant map key information and to display the date of the data being displayed. Additionally the Time Slider updates, displaying a date range under the entire control, highlighting the selected time period of the data in bright yellow, and then using alternating light grey / dark grey boxes to display the temporal extent of other time slices of that dataset. A new time slice can be selected by moving the carrot on the time slider to a new position, which will update the map view and reported date range in the legend. In the example below, the Food Security Conditions layer has three potential time slices.

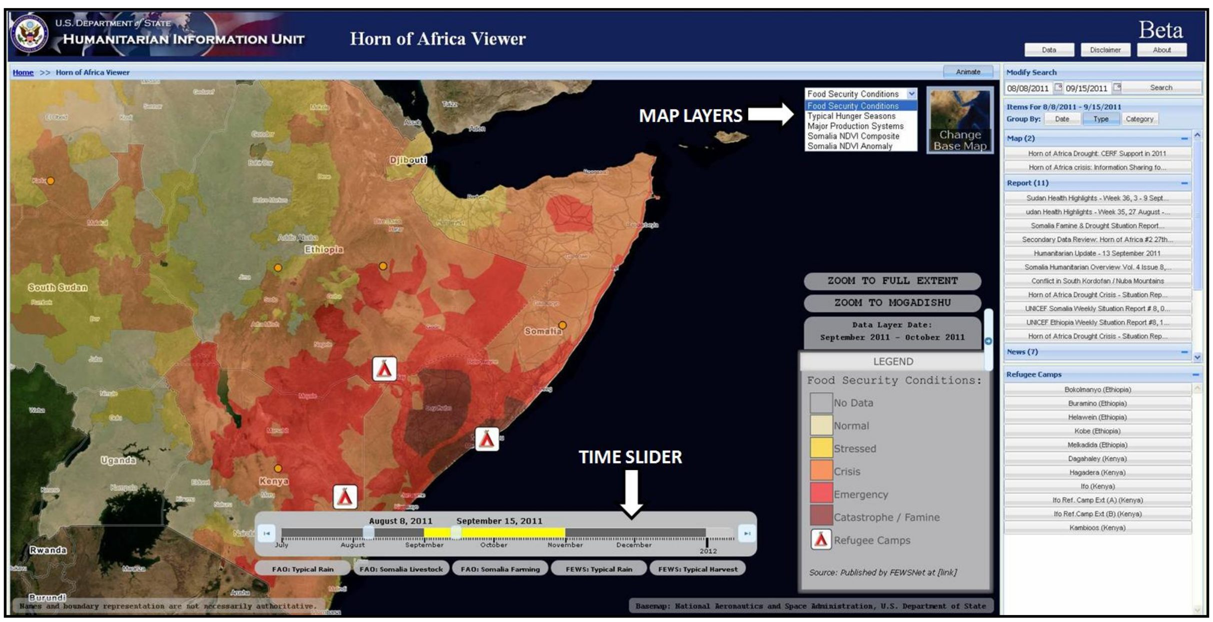

Figure 9 – Screenshot of the HoA Viewer “Time Slider” feature

In order to understand the spatial dimension of the Food Security Conditions, I wanted to compare the reported famine zones to satellite-derived data of vegetation health to understand how much of the famine was due to weather versus caused by humans. We used two vegetation health datasets, both derived from MODIS 16 day NDVI composites. The first dataset is the vegetation health for that 16 day period, and the second a comparison of that specific 16 day period with the 5 year average for that period. As seen in Figure 10 and 11 below, the temporal density of the data is visually discernible in the time slider control.

To further contextualize the vegetation data, we developed the ability to overlay "calendar" feature in the space above the timeline. In this example, the rainy season calendar and agricultural harvest calendar can be toggled on and off and appear in the correct temporal position relative to the time slider. Any type of date range event can be a calendar feature, in this example they are agricultural and climatic events, but in another scenario they could be political or electoral events.

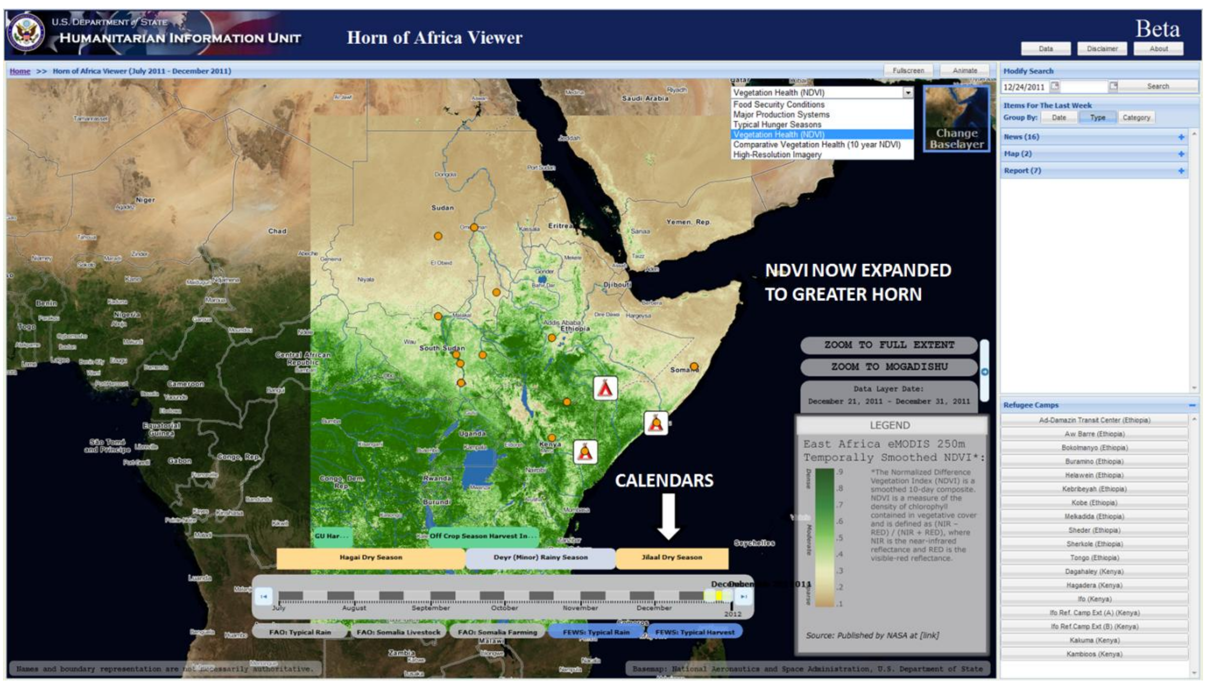

Figure 10 – Screenshot of the HoA Viewer, NDVI data and "Calendar" function. Note the light grey / dark grey bars in the Time Slider control indicating temporal frequency of the dataset.

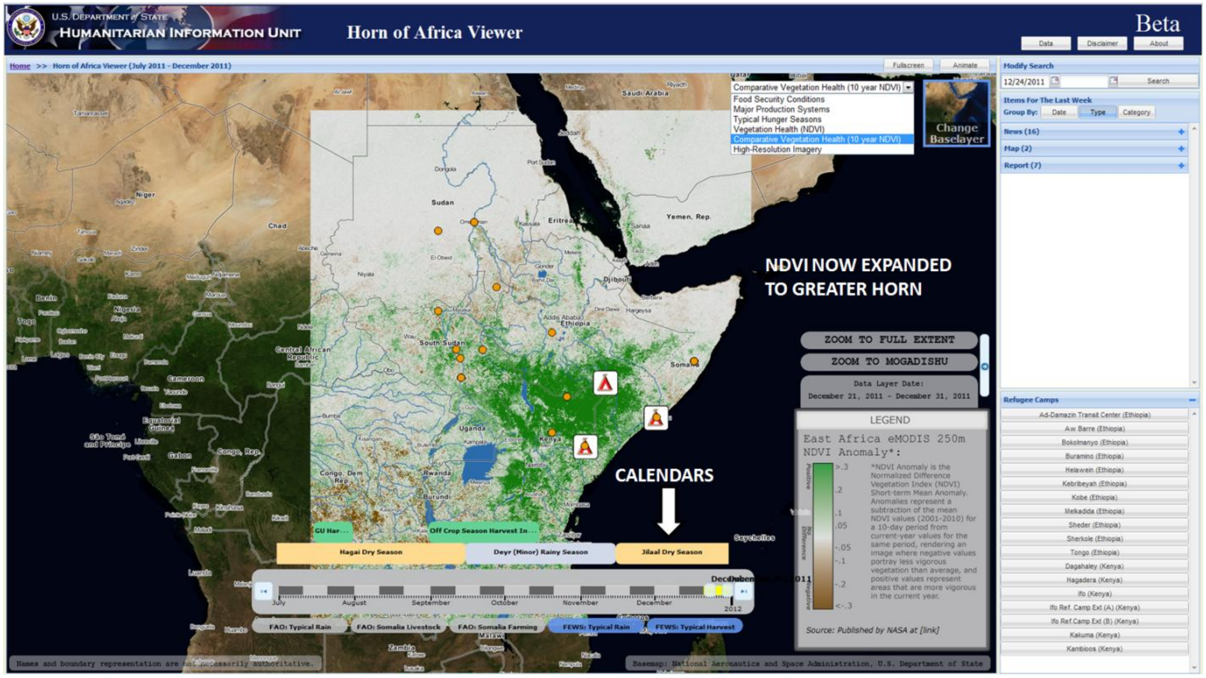

Figure 11 -- Screenshot of the HoA Viewer, NDVI difference for 10 year average data and "Calendar" function. Note the light grey / dark grey bars in the Time Slider control indicating temporal frequency of the dataset.

Curated news events, recorded in the DHNB, were also integrated into the HoA Viewer, as were the search interfaces from the DHNB application. A given news event is geocoded to the closest relevant geographic entity in the story, appearing as a color coded circle icon. On clicked the icon will open a dialogue box displaying relevant information about the news story and a link to the original source. The view below, in Figure 12, displays these elements as well as the fact that the Time Slider control will automatically collapse into a smaller object when not used. Moving the mouse cursor over the box will expand it again.

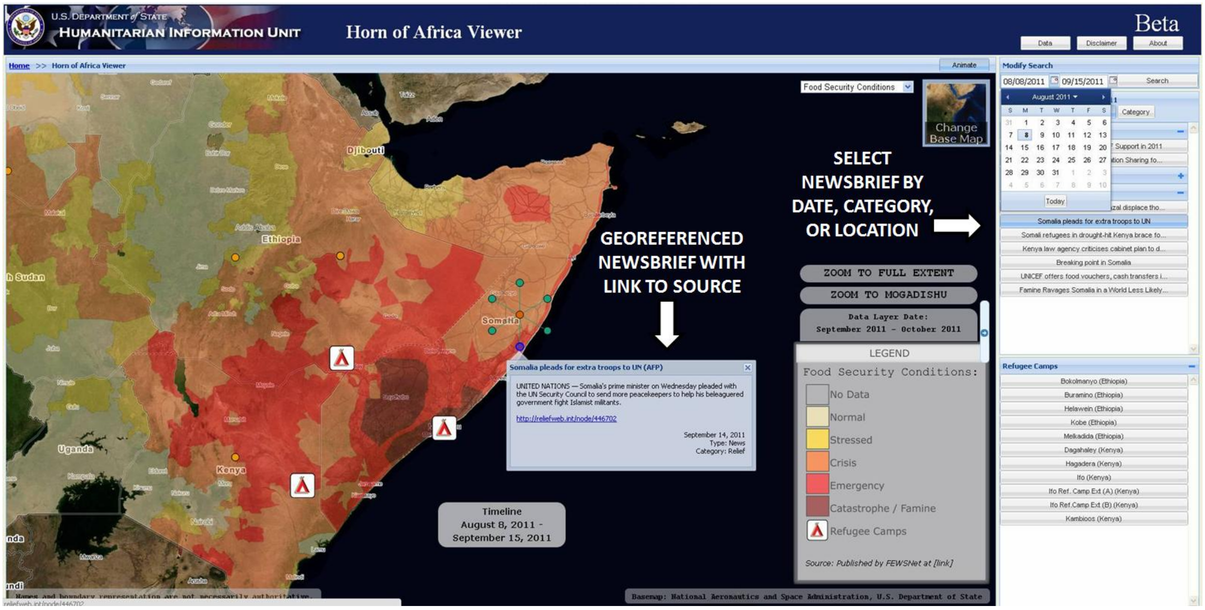

Figure 12 – Screenshot of HoA Viewer, integrated Daily Humanitarian News Brief data feed and user interface controls

The most powerful dataset in the HoA Viewer is the location of large refugee camps in Ethiopia and Kenya and daily update statistics about the camp populations and demographics. The source data on refugees was openly published by the United Nations High Commission on Refugees (UNHCR) through their website. The problem was the data was released in pdf format as numerical tables; HIU researchers manually entered these numbers in a custom HIU database built for the HoA Viewer. The camp data can be accessed by clicking on the map icon or the camp name in the right hand dialogue box. Once selected a dialogue box will open over the camp displaying a snapshot of the current statistics, see Figure 13. This interface was carefully designed to provide the maximum amount of information in the easiest to consume visual format. The Time Slider metaphor is translated from the larger application and appears as a smaller slider at the bottom of the dialogue box. In this case the temporal density of the data is represented by the individual black lines on the slider. The dialogue box opens on the closest data point to whatever date is currently selected in the larger Time Slider, from there the user can explore the refugee data independently. A link to the original UNHCR report for that day is also displayed above the time slider; the link updates depending upon the day selected.

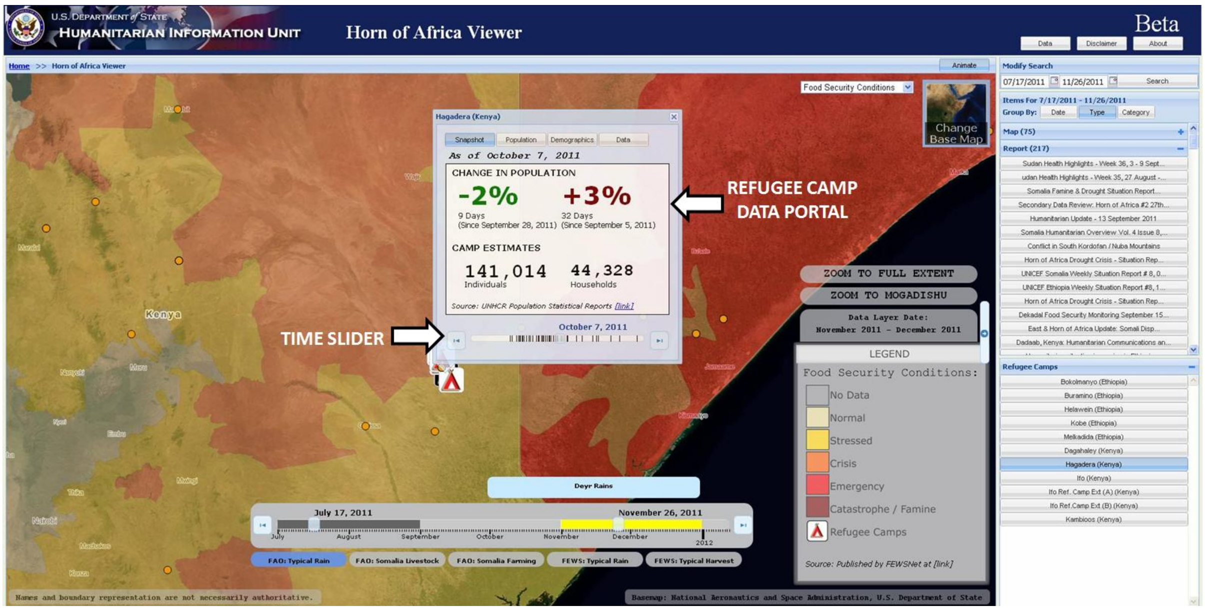

Figure 13 – Screenshot of HoA Viewer, Refugee camp data portal

The initial view of the camp data displays changes in the camp population in the last 7 and 30 days. These values are color-coded to show whether the camp expanded (red) or decrease (green) in size. Below these relative measures are the actual counts of the population, both by individuals and households. Tabs at the top of the dialogue box provide different graphical representations of the camp population data, including trend lines over time (Figure 14), and time-enabled bar charts of the several demographic measures of the camp population (gender, age) (Figure 15). The last tab at the top provides a link to the downloadable HIU version of the refugee population database, available in both proprietary and open data formats (Humanitarian Information Unit 2012).

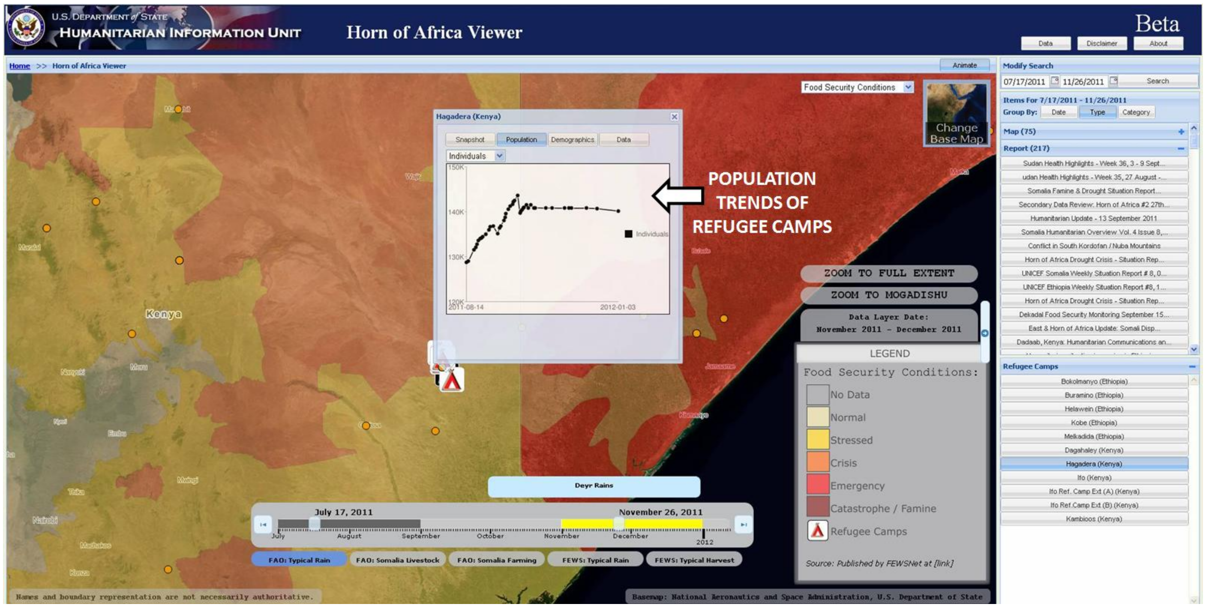

Figure 14 – Screenshot of HoA Viewer, graphical view of refugee camp population data

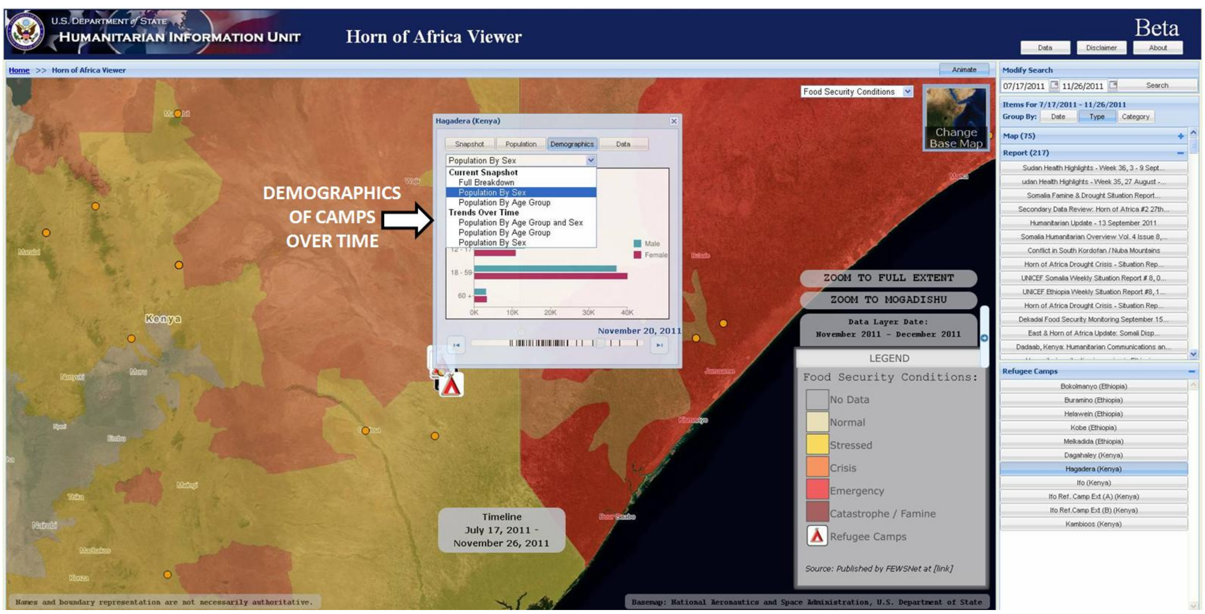

Figure 15 -- Screenshot of HoA Viewer, graphical view of refugee camp demographic data

Lastly, the OpenStreetMap basemap is an option for display as well. Figure 16 below shows the density of data in OSM over Mogadishu, Somalia, the result of the Humanitarian OpenStreetMap team importing a United Nations road database into OSM. Figure 17 also displays the OSM database, this time for the area over the Bokolmanyo refugee camp in Ethiopia. This camp contained approximately 40,000 refugees at that time and was not represented at all in the database. And it was not alone, almost all of the 10 large camps tracked in the HoA Viewer had little to no spatial data in OSM. This data gap provided the first opportunity to test the Imagery to the Crowd methodology, which was being developed at the same time as this application.

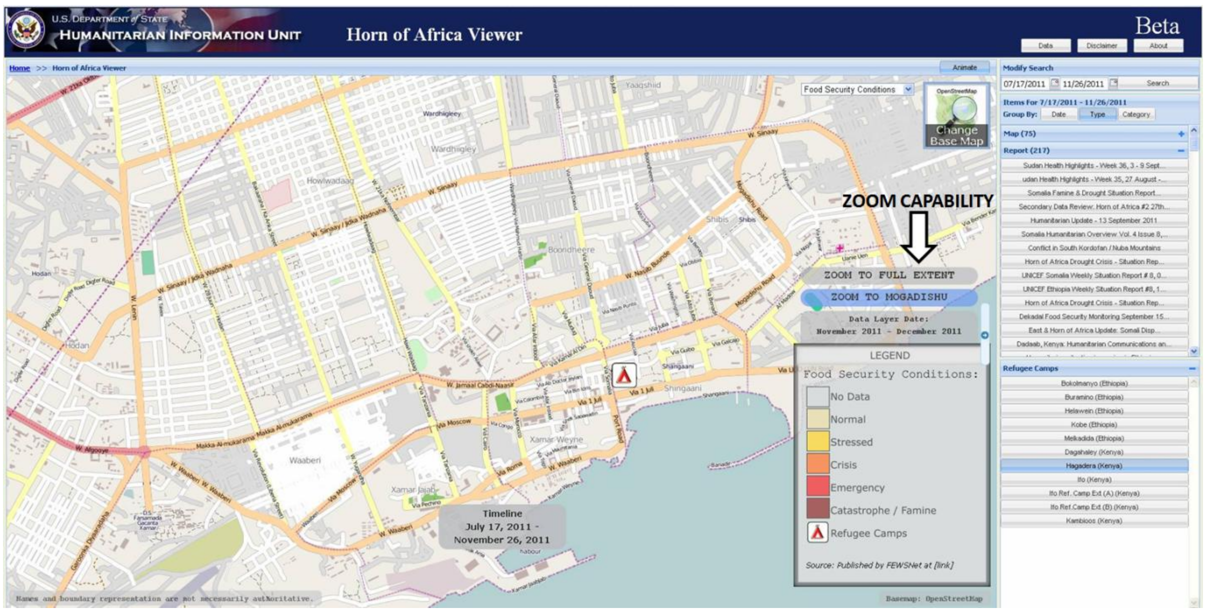

Figure 16 -- Screenshot of HoA Viewer, OpenStreetMap basemap integration

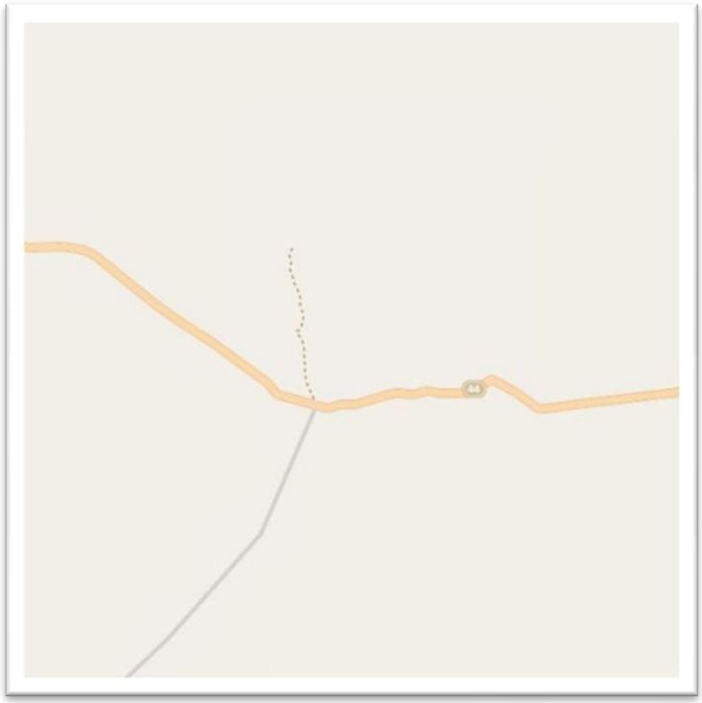

Figure 17 – Screenshot of the Bokolmanyo refugee camp as depicted in the OpenStreetMap database (May 20, 2012)

GeoNode

The development of the CyberGIS spatial data catalog application, GeoNode, was intentionally pursued after the development of the other viewer technologies. As a result its deployment occurred a couple months after my departure from the HIU. That said, GeoNode recently received security approval to be deployed on the OpenNet internal network, and has already been deployed on the HIU's cloud infrastructure. Using the cloud deployment, the HIU released a GeoNode specifically tailored for the Ebola response (ebolageonode.org), and was able to work with a coalition of responding organizations including the Red Cross, United Nations, and World Bank to create a common data repository for the response (Figure 18). While this may not sound impressive, it has historically proven to be very challenging. The reason it was able to happen now is simply one of infrastructure, the HIU had the internal capability and cloud infrastructure.

Figure 18 – Screenshot of the Ebola GeoNode

Crisis Mappers

Starting in 2009 a new community of technologists and humanitarian professionals began to coalesce around the notion of "Crisis Mapping". What started as a small conference of 80 people in Cleveland, Ohio has grown into an international distributed network of thousands of volunteers and humanitarian professionals all working together to bring the power of peer production networks into humanitarian response (Crisis Mappers 2015). The emergence of this discipline is reviewed by Ziemke (2012) describing the central elements of crisis mapping: crowdsourced event data, mapping, imagery, visualization, and analytics.

For the purposes of this report, the Crisis Mapping network is a quintessential model of the Open Source Paradigm Shift applied to humanitarian response. Utilizing network enabled collaboration, open source software, and a peer production of crisis information, the "crisis mappers" have pioneered a new approach to information collection, processing, and analysis (Meier 2012; Liu and Ziemke 2013; Liu 2014). The hallmark of crisis mapping is a workflow that combines crowdsourced event data, typically from social media and SMS / text message reporting, with an interactive map to present a geographic view of rapidly changing events on the ground. The open source software package, Ushahidi, provides the tools and workflow to process unstructured crowdsourced reporting into structured, quantized, translated, and georeferenced entities that are often visualized on an OpenStreetMap basemap (Figure 19)(Okolloh 2009).

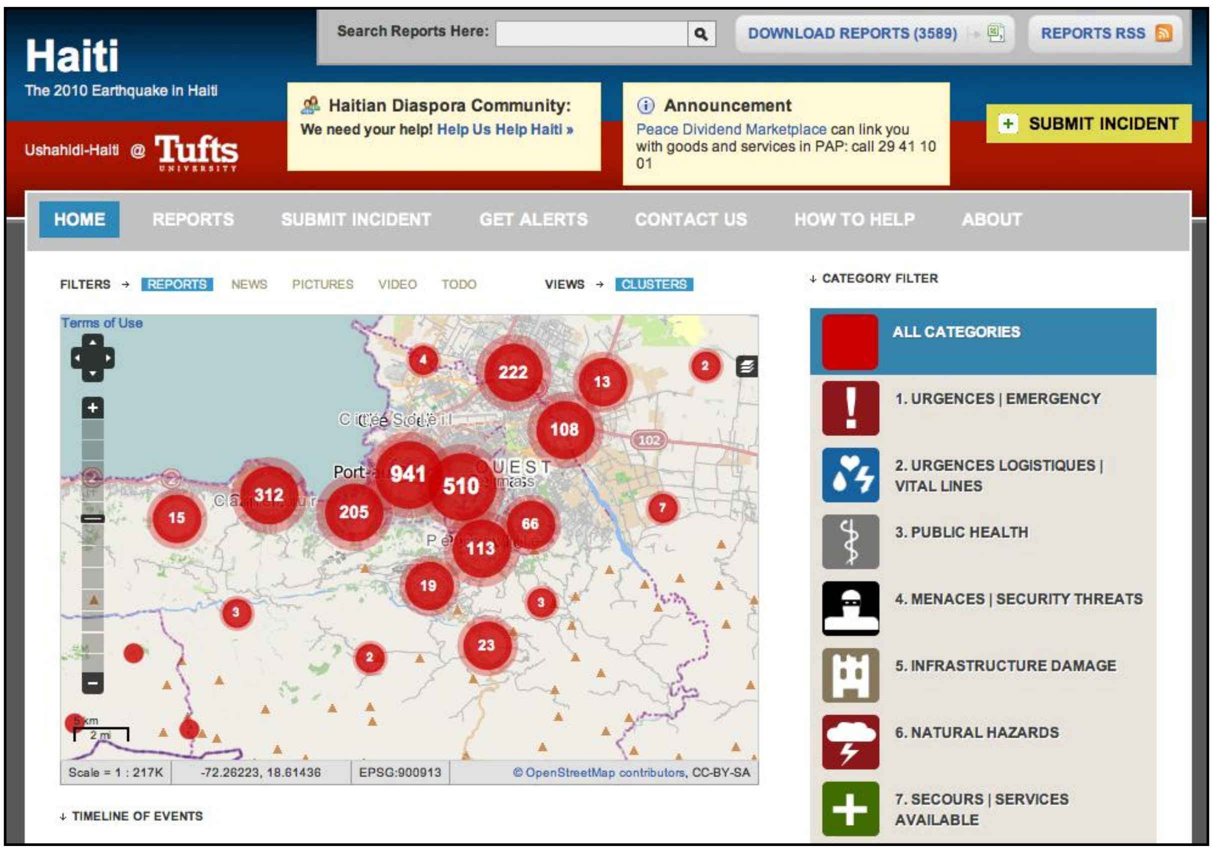

Figure 19 – Screenshot of Ushahidi Haiti application

Over the last five years this loose affiliation of volunteers has sub-divided into many different peer production volunteer networks specializing in particular skills, including the Standby Task Force (information processing), Humanitarian OpenStreetMap Team (mapping), DataKind (data manipulation), etc. In conjunction with the United Nations Office for the Coordination of Humanitarian Affairs (UN OCHA) these various volunteer groups formalized their activation procedures in an effort to formally engage with the traditional humanitarian community. This process resulted in the formation of the Digital Humanitarian Network and an established methodology by with the humanitarian community can formally request help from volunteer organizations, with the expectation that help will come (Figure 20)(Capelo, Chang, and Verity 2012). Just as the peer production networks that created Linux and OpenStreetMap, the Crisis Mappers have proven their utility and become an indispensable element of humanitarian response.

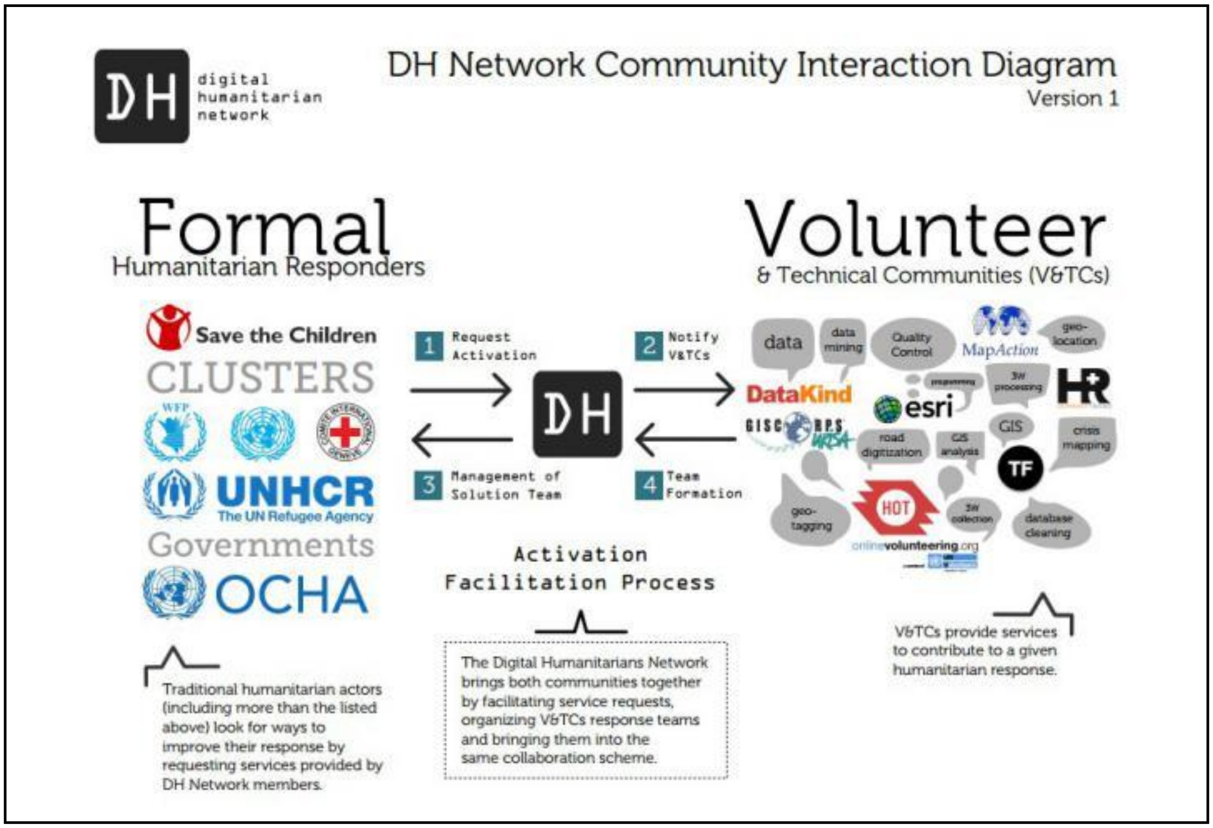

Figure 20 – Digital Humanitarian Network activation process

I'm proud to have been a part of the original crisis mappers, having attended the first International Conference on Crisis Mapping, and am one of only three people who've attended all five. The first event in Cleveland started a revolution, and we all knew it. I was struck at the time how much the meeting seemed like a case study of "lead users", a concept devised by von Hippel (1986) for the creation of novel product concepts. My role in this organization has always been to promote the role of geographic data in humanitarian response. As stated in an earlier section, the seeming ubiquity of Google Maps often lulls people into thinking dense geographic base data is available everywhere, when it isn't. Social media reporting is a powerful tool that can help direct relief efforts in real time, but only if those reports can be geocoded and visualized.

Haiti Earthquake Response

The first Crisis Mappers conference was held in October 2009, and as stated above, there was a distinct feeling that crisis mapping was going to have a major impact on disaster response. Unfortunately, I did not take long to test it. The Haiti earthquake struck on January 12, 2010, and that response introduced crisis mapping to the world. The role of crisis mapping has been detailed in several reports and articles, the best overall perspective from the humanitarian community is Crowley and Chan (2010) Disaster Response 2.0, Meier and Munro (2010) review crowdsourced reporting using SMS, and Zook et al. (2010) discuss the role of crowdsourced mapping and volunteered geographic information.

It is the role OSM mapping played in the response, and the methods of how it was produced that are useful here. Haiti was the first significant activation of the Humanitarian OpenStreetMap Team. Almost immediately after the earthquake struck, volunteers started digitizing old street maps and anything they could find to build up OSM in the areas around Port-au-Prince[1]. Figure 21 below is a snapshot of what the city looked like in OSM before the earthquake.

[1] As an aside, immediately after the earthquake I worked with Matt Dunbar, another KU Geography PhD, to publish old basemaps of Haiti and Port-au-Prince as web services that could be used by volunteer mappers; these services are still running on the KARS WebGIS platform (http://kars.ku.edu/research/haiti-earthquake-disaster-portal/)

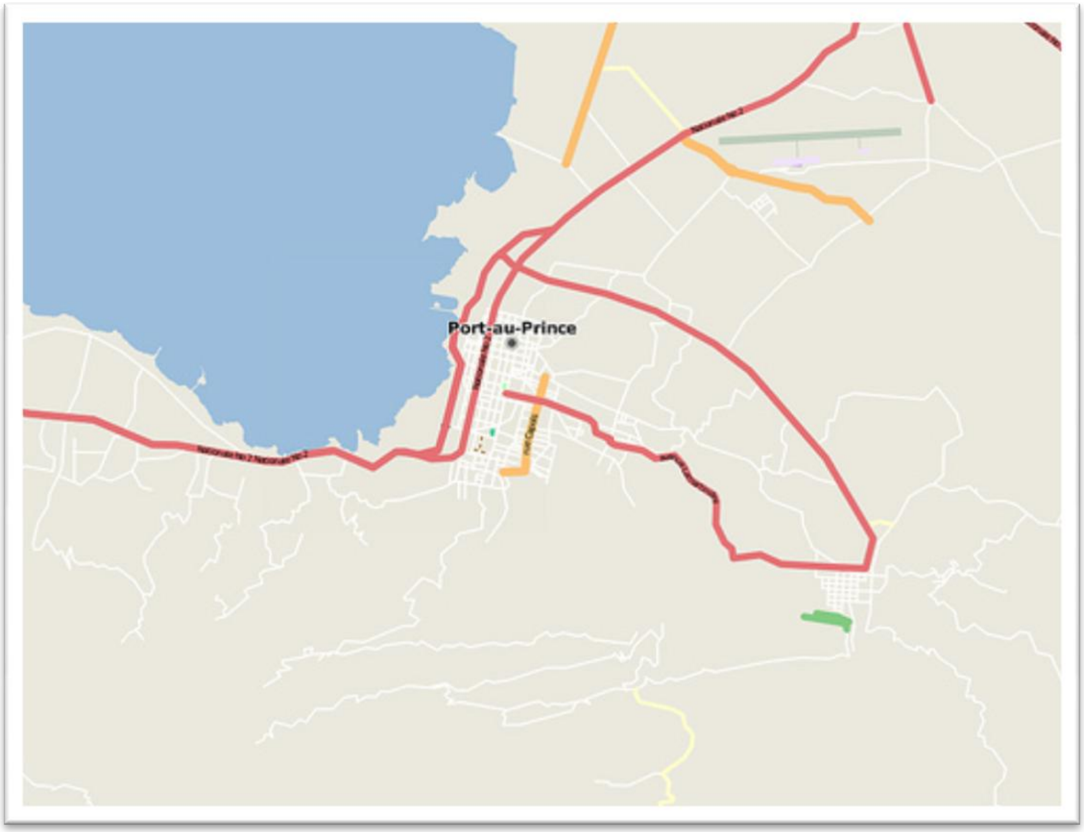

Figure 21 – Port-au-Prince in OSM before January 12, 2010
(http://www.flickr.com/photos/itoworld/4351891226/sizes/m/in/photostream/

Within days after the earthquake, the commercial satellite imagery companies started freely releasing satellite imagery over the affected area under an open license that allowed it to be traced by OSM volunteers. The transformation in the OSM basemap was incredible, see Figure 22 below, and soon the OSM basemap became the de facto operational basemap of the response.

Figure 22 – Port-au-Prince in OSM in February 2010
(http://www.flickr.com/photos/itoworld/4351891526/sizes/m/in/photostream/)

For me, the critical question coming out of this response concerned whether this level of volunteer effort was a one-off event spawned by the magnitude of the disaster, or if it was repeatable? It was clear that the provision of updated satellite imagery provided a catalyst to mapping efforts, but the imagery was massive in size and needed to be processed for easy consumption into the OSM editing tools. In this case a couple of volunteers with a background in imagery and software development had a huge impact. Using server space and bandwidth donated by the University of San Diego Supercomputing Center, they processed all of the available imagery and published it as web services in an application called the Haiti Crisis Map, see Figure 23 below. These web services could also be consumed directly in the OSM editing software.

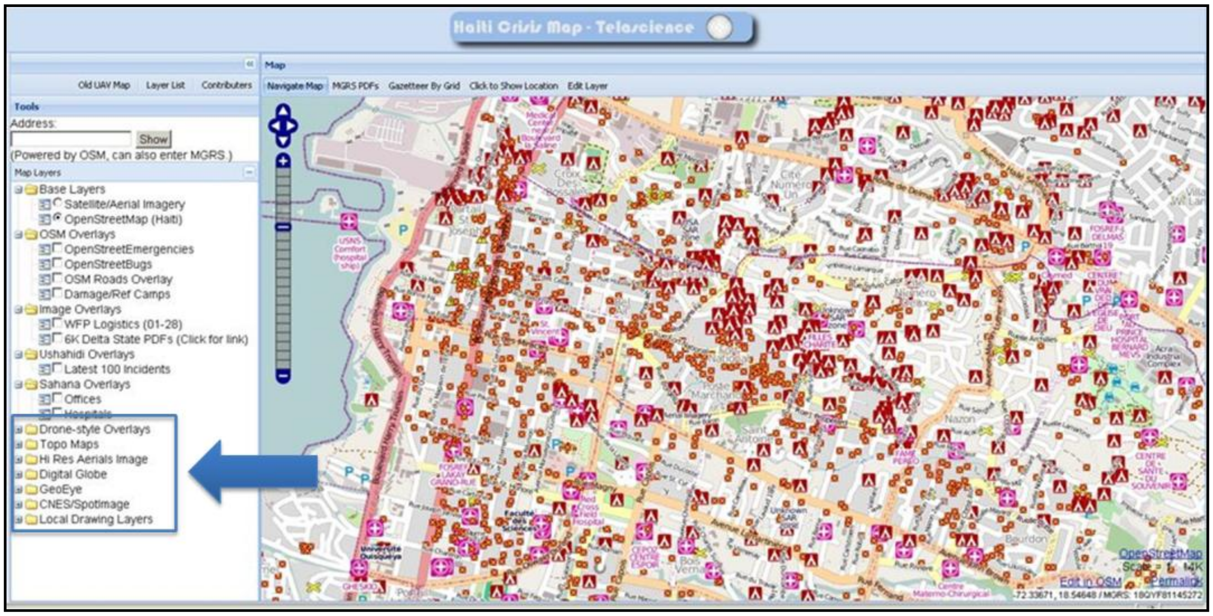

Figure 23 – Screenshot of the Haiti Crisis Map, highlighted area on lower left side display links to the various imagery datasets processed and published by this application

This much image processing was a tremendous effort, but it enabled the broader volunteer community to donate their time mapping and not image processing. Unfortunately, six months later, a system administrator cleaning up server space accidently deleted the entire set of imagery caches, all 15 terabytes, gone in an instant. So in terms of determining if the process was repeatable, the critical elements included updated imagery, dedicated computing and bandwidth resources, experienced personnel who knew imagery and software, and a web service publishing model that allowed volunteers to easily connect to the imagery. This seemed to be the recipe for supporting and catalyzing volunteer mapping effort, so the next questions became, could the HIU provide all of those requirements? And would the crowd come help?

Imagery to the Crowd

Born from the conceptual framework of GIS 2.0 and inspired by the volunteer mapping response to the Haiti earthquake, Imagery to the Crowd (IttC) was designed to be a repeatable, sustainable mechanism for the U.S. Government to help catalyze and direct volunteer mapping efforts. The goal is to produce geographic data about areas at risk of, or experiencing a complex emergency. These situations range from fast onset natural disasters, typhoons, floods, epidemiological outbreaks, to slow onset issues of food security or political and economic security issues that lead to large scale human migration, refugees and internally displaced people. Beyond being just reactive to disasters, the intent was to also identify areas at risk, and working with partners, help produce data for disaster risk reduction and community resilience projects.

IttC (pronounced *itsy*) is an intentional attempt to link the cognitive surplus of the crowd, in the form of the volunteer network behind the Humanitarian OpenStreetMap Team (HOT), with the purchasing power of the Unites States Government, to catalyze the production of quality, open geographic foundation data. As was reviewed in the Haiti example, it was the release of current, high resolution satellite imagery published as web services that enabled the crowd to focus on mapping. In reviewing the critical elements of what made the Haiti mapping a success, it was the availability of current imagery, the computing infrastructure to publish it, and the personnel with the right combination of skills that made it possible. So the key question moving beyond Haiti was whether Humanitarian Information Unit could fill all these roles and help catalyze the crowd.

IttC was predicated on the belief that the USG could provide the key structural elements need to support volunteer mapping. These baseline requirements included:

1. Vetted humanitarian need to assure that volunteer efforts directly supported humanitarian and resilience organizations that would immediately benefit from the availability of updated data
2. Updated commercial high resolution satellite imagery
3. Imagery formatted as web services that could be consumed in OpenStreetMap editing software
4. Process to ensure that data created by volunteers would be stored in OpenStreetMap, and not controlled by the USG

To accomplish this, several components were necessary. First, we had to determine if the USG had the right within its existing contractual structure to share imagery-based web services with the then nascent crisis mapping community. This portion of IttC dealt with the NextView license, the agreement that defined the sharing provisions for imagery purchased through the NextView contract vehicle. Second, we had to build a geographic computing infrastructure and organizational capability to ingest, process, and publish imagery-based web services at a speed and scale to be useful. Accomplishing this was a significant technical challenge, and leveraged the HIU CyberGIS. Third, we had to gain the trust of the crisis mapping community that the U.S. Government was actually "here to help."

Understanding how Imagery to the Crowd came into existence requires some background on the history of satellite imagery sharing in the USG. Issues related to the sharing of high resolution satellite imagery between the USG and partner organizations goes back to the early 2000s when the first USG commercial imagery contracts were established. Specifically it is the sharing agreement built into the NextView contract that is central to Imagery to the Crowd. However, because volunteer mapping was such a new concept, IttC presented a new set of issues than previously encountered and required a separate review. The full story of satellite imagery sharing is beyond the scope of this report, but this effort built upon extensive efforts of the Bureau of Intelligence and Research, Office of the Geographer and the National Geospatial-Intelligence Agency. In terms of sharing NextView imagery with the humanitarian community, significant work was also done in 2009 for projects in Afghanistan (Crowley 2015). Given the long history and multiple groups involved, the intent here is to only include the elements of the NextView history, timeline, and conclusions that relate to IttC and that I had direct involvement with.

NextView / EnhancedView Contracts

Beginning in 2003, the USG, led by NGA, established contract vehicles to procure commercially available, high resolution satellite imagery from the firms DigitalGlobe and GeoEye. The first of these contracts was titled "ClearView", its replacement was called "NextView", and since 2010 the latest version is called "EnhancedView". The timing and dynamics of these contracts varied in the early years, but beginning in 2005 DigitalGlobe and NGA agreed to terms on the USG's ability to share imagery purchased under the NextView contract, a document that is referred to as the NextView license. By 2007, both DigitalGlobe and GeoEye were bound by the sharing provisions in the NextView license, and since 2012 when DigitalGlobe acquired GeoEye, all commercial imagery purchased by the USG is licensed under the NextView license. With the transition to the EnhancedView contract vehicle in 2010, the NextView license was carried over and is still in effect.

Since 2007, the NextView and EnhancedView contracts have been structured as multiple year agreements with renewable option years annually. EnhancedView started in 2010 and lasts until

2019. There are several components to the contract, but the core is the Service Level Agreement (SLA) that provides imagery services to NGA, essentially a baseline amount of imagery per day plus a certain amount of tasking priority for a set monthly price. With the launch of the WorldView-3 satellite, the current payment for that SLA is $25M per month, for a total of $300M annually (DigitalGlobe, Inc. 2015). The amount of imagery and tasking are both redacted in the SEC 10-K filing cited above, so the specifics of the contract are not publically reported, not surprising given NGA's membership in the Intelligence Community. This same report indicates that 60.4% of DigitalGlobe's 2014 revenue ($654.6M) comes from the USG, and 55.3% from the EnhancedView contract. This is simply to demonstrate that the USG is the key customer for the company. Some additional characteristics of the EnhancedView contract, as well as some history, is reviewed in Crampton et al. (2014).

NextView License

The NextView license is the key component in the Imagery to the Crowd and MapGive projects. It is the legal agreement between DigitalGlobe and the United States Government that defines the sharing provisions for commercial imagery purchased under the NextView and EnhancedView contracts (see Appendix 1 for complete license). A digital copy of the license is included with every image that is downloaded or delivered to the USG.

In my opinion, the NextView license is an incredible document in terms of the latitude of sharing permission that it grants the USG, and it is a credit to the government officials and contracting officers that negotiated it. The EnhancedView contract represents a significant investment by the USG in satellite imagery, and the NextView license ensures that we as tax-payers can get the most benefit from that investment. The license is a brief 1-page document, one that I believe is fairly straight forward to interpret. However, for a multitude of reasons, it has proven difficult to determine exactly what is allowable under the license, and therefore, to extract the value from the license that we are paying for. To understand how IttC evolved, we have to start with the exact language of the license. What follows is an annotated version of the license, with interpretations of the key elements.

> **<u>NEXTVIEW IMAGERY END USER LICENSE AGREEMENT</u>**
>
> 1. Introduction. This End User License Agreement ("EULA") is between DigitalGlobe, Inc., a Delaware Corporation ("DigitalGlobe" or "Seller") and National Geospatial-Intelligence Agency ("NGA"), the purchaser of this EULA, which governs the use of the data products or documentation ("Products") accompanying this EULA in accordance with Contract NMA 301-03-3-0001 (the "Contract").
>
> 2. Applicability. This license applies to imagery and products licensed under the Contract, including data downlinked to domestic and foreign ground stations.

The above clauses defines this document as the *End User License Agreement (EULA)* and further establishes the two parties in the contract, the contract vehicle, the products that the EULA covers, the generic terms that will be used, and the geographic extent.

> 3. License Granted and Permitted Uses.
>
> a. General Terms
>
>> 1. This clause applies to all unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data licensed under this Contract. No other clauses related to intellectual property or data rights of any sort shall have any effect related to the unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data delivered under this Contract.

This first clause under the General Terms heading establishes two things, first, that this EULA applies to all *unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data* that are delivered as part of the

Contract. The key element for IttC is the “requirements-compliant processed imagery” phrase, as most imagery used in the project has already been processed to some degree. Second, that there are *No other clauses…shall have any effect related to the …data delivered under this contract*, essentially saying there is no other document that will supersede the terms established in this document. This clause is critical because it expressly states that the EULA is the governing document for all products purchased under NextView and now EnhancedView.

> 2. All license rights for use of the unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data provided to the U.S. Government purchased under this NGA contract are in perpetuity.

This clause states that the rights granted under this EULA apply to the data purchased *in perpetuity*, another key element as it means products purchased at one point during the Contract retain these terms into the future.

> 3. Licensed users may generate an unlimited number of hardcopies and softcopies of the unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data for their use.

This clause states that *Licensed Users*, who are defined in the following clauses, have the right to make unlimited copies (both hardcopy and softcopy, *i.e.*, digital reproduction) of Products purchased under the Contract.

> 4. (i) Licensed users may generate any derived product from the licensed unprocessed sensor data; and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data.

This clause states that a Licensed user can modify the imagery to make *any derived product*. This places no limitation on the amount or type of analysis or dissemination formats.

(ii) Unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data licensed under this NGA contract have no restrictions on use and distribution, but shall contain the copyright markings.

This clause is again key to IttC, it expressly states that Products purchased under this Contract have *no restrictions on use and distribution*, with the only caveat that the *copyright markings* be maintained on any distributed product.

b. Licensed Users

1. The imagery may be used by the U.S. Government (including, all branches, departments, agencies, and offices).

2. The U.S. Government may provide the imagery to the following organizations:
State Governments
Local Governments
Foreign Governments and inter-governmental organizations
NGO's and other non-profit organizations

This section defines two things related to *Licensed Users*. First, in clause 1, this Contract is intended to be accessible to the entire U.S. Government. So while NGA is responsible for maintaining the contract, they do not act as the gate keeper or adjudicator of who uses the imagery and for what purpose. This clause stands in agreement with NGA's position as a "support agency", meaning their mandate is to support other Lead Federal Agencies, who have congressionally granted authority over particular elements of the government (Department of Defense for defense, State Department for foreign policy, Department of Interior for management of domestic natural resources, etc…). Second, clause 2 clearly states that the U.S. Government may provide imagery to a range of other organizations, including State, Local, and Foreign governments, inter-governmental organizations (examples include the United Nations, NATO, and World Bank), and non-governmental organizations (NGOs) and other non-profit organizations. Clearly the latitude of clause 2 is to ensure that the USG can share the imagery it purchases with almost any organization that it needs to.

> 3. In consideration for the flexibility afforded to the U.S. Government by allowing unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data to be shared, the United States Government shall use its reasonable best efforts to minimize the effects on commercial sales. Acquisition and dissemination of imagery and imagery products collected within the United States shall be restricted in accordance with law and regulation.

This is the only clause that places any constraint on the sharing of Products under the Contract. Two phrases are key here, *In consideration for the flexibility afforded to the U.S. Government*, a clear nod to the fact that the EULA is very liberal in granting USG broad rights on sharing imagery, and *the United States Government will use its reasonable best efforts to minimize the effects on commercial sales*. It is this latter statement that is the key element that had to be interpreted for IttC, and more broadly, it is this clause that leads to variability in the interpretation of the EULA. Several questions emerge from this clause:

- How do you define what a reasonable best effort is?
- What is a commercial sale? Only to commercial / for-profit entities?
- Are inter-governmental organizations or NGOs that might purchase imagery considered a commercial sale?
- What does minimize mean?

As will be discussed below, there are other guidance documents provided by NGA that designed to explain the guidelines for sharing imagery. However, these documents can be contradictory and / or less precise than the EULA. This makes it difficult to determine what exactly is permissible, and worse, potentially violates the EULA as General Terms, Clause 1, states there is no document that supersedes the term of the EULA.

> DigitalGlobe, Inc.
> NextView EULA_5749 Rev 1.0
> 08/10/05

Finally, this closing element is worth noting as it is dated to August 10, 2005, and is still in effect today.

NextView and You

In order to provide standard guidance on how to interpret the NextView license, NGA has provided a supplementary document entitled "NextView License and You" (see Figure 24 below and Appendix 4). For the most part, it does simplify the NextView EULA, but also introduces definitions of product types that do not exist in the EULA and some caveats about dissemination for those product types. Most problematic is the introduction of a false dichotomy between "imagery" and "imagery derived products (IDP)". On one hand it treats the two similarly, as in

> You may share ***imagery or IDPs*** with anyone directly working with/for the USG, including …
>
> ***U.S. Government Employees / Contractors***,
>
> ***Universities supporting USG via contract(s),***
>
> State / Local Governments,
>
> Foreign Governments,
>
> Intergovernmental Agencies,
>
> NGOs & Non-profit Organizations

Note the additional interpreted distinction of the original Licensed Users to include USG employees and contractors, and Universities working for USG via contract or grant. However, in the next two bullet points it states,

> You may:
>
> Post properly attributed "dumb" IDPs on public web sites
>
> Post/disseminate imagery using access-controlled web/FTE sites" (emphasis in the original document, assume FTE is a typo and they meant FTP site)

It does not make sense that you are simultaneously limited in who you can share an IDP with and be able to post it on an open website.

In the EULA, General Terms, Clause 4(ii) states there is no restriction on use or distribution of *unprocessed sensor data, requirements-compliant processed imagery, imagery services, imagery derived products and imagery support data, imagery or imagery-derived products*, yet here, a distinction between the allowable modes of dissemination of an IDP and Imagery has been introduced, mainly what can go on a "public" website versus an "access-controlled" website. This distinction is somewhat understandable, as it is trying to take into consideration the commercial clause and the fact that the USG should not just be giving away imagery. However, this is an important distinction as it affects a formatting decision that IttC has to make.

With these two rules now in place, the question turns to the distinction between the Imagery and Imagery-Derived Product types. The "Definitions" section at the bottom of the document provides a description of the key terms and distinctions, including Imagery, Imagery Derived Product, and Non-literal Imagery Derived Product:

Imagery is the image and associated metadata. Imagery can be further manipulated, enhanced, & processed. Example: GeoPDF, GeoTiff, NITF.

Image Derived Product (IDP) – any product created from raw imagery – could include image metadata, but generally does not and often referred to as "dumb-image"

Non-literal Imagery Derived Product is a product derived from the imagery but no longer looks like an image. Example: Line-drawings, maps.

From a remote sensing perspective, these definitions are lacking clarity in that they conflate the properties of raw imagery with the potential formats it could be encoded in, and beyond that, the formats selected do not consider the implications of web services and how they deliver geographic data over the network. Recommendations for changing these definitions are discussed in more detail in a later section.

Figure 24 – The "NextView License and You" document

NextView License Acknowledgement

There is one additional NextView document that is relevant to IttC, the License Acknowledgement. This is a paragraph-long document that accompanies any imagery that is shared with partners (see Appendix 3). Essentially the Acknowledgement is to ensure that a user understands they are accepting the terms of the NextView License, a copy of which is also included. The key element to this document is the final line, "Your acceptance of these license terms is implied by your use." Therefore, this acknowledgment document does two things, informs the user of the NextView license and some of its terms, and obligates the user to follow the NextView terms if they use the imagery. The latter is needed to demonstrate the USG is doing its part to ensure that partners respect the NextView terms.

Existing Imagery Sharing Protocols

It is important to recognize that the Department of State, mainly through the Humanitarian Information Unit, has been sharing imagery with a range of partners from the beginning of the ClearView/NextView contracts. There is an established internal process for vetting requests that includes a policy review of the operational partners and the justification for the release. This process requires that the relevant elements at State Department and USAID have to approve that the release of imagery to a specific group, for a specific purpose, is in the USG "best interests". This process ensures that there is a high level of support within the State Department and USAID for a particular imagery release. This review process demonstrates the value placed of making the correct policy decision regarding imagery sharing, as well as a concern to operate within the reasonable best efforts constraint of the EULA. However, the NextView EULA does not require this level of review, and any NextView considerations have come from additional interpreted guidance. Once a request is approved, imagery is either shipped on disc to the partner or pushed from the online imagery archives to the recipients FTP site.

IttC Workflow

The Imagery to the Crowd (IttC) project took two years to develop, and required that several legal, policy, technology, and organizational challenges be addressed before it could be implemented. In the following sections each of these factors is discussed, told in the order they were addressed, concluding with the launch of the IttC prototype and the results from some of the notable IttC projects.

IttC had its genesis in a series of meetings that began in October 2010 around the 2nd International Conference on Crisis Mapping (ICCM 2010) between John Crowley (National Defense University & Humanitarian OpenStreetMap Team), Dick Williams (NGA), and Joshua Campbell (HIU). This initial meeting led to a series of follow-on meetings between NGA, State Department, and individuals from the broader crisis mapping community. The first of the follow-on meetings began with a December 2010 meeting held at the HIU, and again in February 2011 at the Research & Experimentation for Local & International Emergency & First Responders (RELIEF) 11-2 workshop in DC. The topic for all of these meetings was how NGA and the broader National System for Geospatial-Intelligence (NSG)[2] could interface with the emerging "volunteer technical community (VTC)", primarily in the context of disaster response. "Volunteer & Technical Community" was the name given to the nascent crisis mapping community in the Disaster Relief 2.0 report detailing the Haiti response (Crowley and Chan 2010), and while it is seldom used now, it was the term of art used to describe the crisis mapping community in the early imagery sharing discussions.

Coming out the RELIEF 11-2 meeting, the thought was that NGA could best explore the options for how to engage with the VTC through a formal experiment, also called a table-top exercise.

---

[2] The National System for Geospatial Intelligence (NSG) is defined in a 2007 report as "the combination of technology, policies, capabilities, doctrine, activities, people, data, and communities necessary to produce geospatial intelligence (GEOINT) in an integrated multi-intelligence, multi-domain environment. The NSG includes the Intelligence Community (IC), the Joint Staff, the Military Departments (to include the Services), the Combatant Commands (COCOMs), international partners, National Applications Office, Civil Applications Committee members, industry, academia, Defense service providers, and civil community service providers." (https://www.hsdl.org/?view&did=19363)

So planning began for the construction of a Limited Objective Experiment with the VTC to address four key themes: Legal/Policy Review, Technical Interoperability (tabletop exercise), Awareness and Familiarization with the Volunteer Technical Community (VTC), and a workflow for engaging in future efforts (Naval Postgraduate School 2011). Planning took place over a series of conference calls, and the experiment was scheduled for the RELIEF 11-4 exercises to be held in August 2011.

It is worth noting that nearly all of the early IttC meetings were held as part of RELIEF experiments. RELIEF was created when humanitarian researchers at the Center for Technology and National Security Policy at the National Defense University began to integrate with an on-going set of experiments held by the Naval Postgraduate School and U.S. Special Operations Command. The decision to work the imagery sharing process through the RELIEF experiments was a significant benefit, as the experiments were designed to be "conducted in a multi-institutional field setting; providing a semi-structured learning environment capable of promoting collaboration and relationship building across an increasingly diverse governmental and civilian response network" (Oros 2014, 16). At that point, John Crowley, working on contract to the National Defense University, was a program lead for the RELIEF project and was playing a pivotal role in coordinating with NGA and the crisis mapping community. RELIEF events were held quarterly, with most located at Camp Roberts, a National Guard base located in central California near Paso Robles, and usually one a year held in Washington, DC. The goal of RELIEF is to create an environment where the complex technological, organizational, and bureaucratic challenges related to disaster relief, stabilization, and reconstruction could be addressed. RELIEF's hallmark was the ability to bring together the formal and informal networks (humanitarian practitioners, technologists, government and military) for hands-on collaboration.

As the experiment planning for RELIEF 11-4 was ramping up, a process conducted primarily by NDU and NGA, I had been working the details of how the imagery sharing process could work within the HIU. Having studied the NextView EULA and License Acknowledgement, the NextView License and You document, and the existing State Department procedure for sharing imagery, I believed that IttC was legally possible, but that the ability to move forward hinged on a clear definition on several key factors:

- Recognition of the State Department as the Lead Federal Agency for imagery sharing related to foreign policy, including complex emergencies
- Determine whether the crisis mapping community could qualify as a sharing partner
- Determine whether a Tiled Map Service met the requirements of an Imagery Derived Product
- And most importantly, determine whether the extracted vectors, mapped from NextView licensed web services and stored in the OpenStreetMap database, violate the commercial sales clause of the EULA.

As hoped for, the breakthrough effort occurred in August 2011 at the RELIEF 11-4 experiments held at Camp Roberts in California. In the tabletop exercise, a disaster scenario involving a hurricane in the Caribbean was used to work through a process for providing commercial satellite imagery to the VTC. What emerged from the table top exercise was an outline for how imagery sharing could be implemented, this outline later evolved into the Imagery to the Crowd process.

The most important portion of the experiment was a teleconference back to DC with the assembled group at Camp Roberts and members of NGA General Counsel, most notably the lawyer who negotiated the NextView license. The goal of the teleconference was to investigate the legal and policy considerations behind sharing NextView imagery with the VTC. Key considerations in the discussion were the role of the Lead Federal Agency in determining who to share imagery with, NGA's role in supporting the Lead Federal Agency, whether the VTC could be considered a Licensed User, and the legal status of vector features extracted from NextView licensed imagery. The assembled group included members of several NGA Offices, USAID/Office of Foreign Disaster Assistance (OFDA), State Department/Office of the Geographer, Office of Naval Research, and leaders in the VTC.

During that discussion the outlines of a legal framework were defined that reinforced the role of the Lead Federal Agency (USAID OFDA for disasters and DoS for complex emergencies) in determining appropriate partners to work with, that NGA's role was to support the Lead Federal Agencies, and the critical distinction, that vector features extracted from NextView licensed imagery and stored in the OpenStreetMap database would not violate the NextView terms. Getting to this answer was a difficult process, having taken almost a full year. The picture below (Figure 25) was taken with my iPhone just after the NGA General Counsel made the statement on the teleconference – the assembled group's reaction speaks for itself.

Figure 25: Photo taken at RELIEF 11-4 during teleconference regarding sharing NextView licensed imagery with the VTC. I'm seated near the camera in the blue shirt.

Immediately following RELIEF 11-4 work continued on refining the process by which NextView imagery could be obtained by a Lead Federal Agency (LFA) and shared with the VTC for mapping into OpenStreetMap. Building on the legal outlines from the RELIEF teleconference, it was further established there is a difference between "pixels" and "non-pixels", and that the NextView provisions applied only to "pixels". This distinction then meant that imagery-derived products and vector data extracted from imagery were not regulated by NextView. An example of this would be a road that was mapped from NextView licensed imagery. If you wanted to share the raw imagery over that road, then NextView applied. But, if you viewed the imagery in a GIS and mapped the road into a vector line, then that vector would not be impacted by NextView and could be shared openly. Additionally you could make a map product with the raw imagery as a backdrop and the road vector styled on the map, export the map into a graphic file (.jpg, .tiff, etc), and the resulting "imagery-derived product" would not be subject to NextView.

From these legal conclusions, the HIU prototyped two technological options for sharing imagery with the Humanitarian OpenStreetMap Team. At that point, late 2011, the HIU CyberGIS was not complete, so we were trying to work with partners on an imagery hosting plan. The first option, for HOT to host the imagery, was quickly ruled out due to limitations in staffing and resourcing. However, we did receive significant technological guidance and support from HOT members during this time, which significantly improved elements of the final imagery sharing plan. The second option was to work with the United Nations Institute for Training and Research Operational Satellite Applications Programme (UNOSAT), whom the HIU already had an existing imagery sharing relationship and who was building a web mapping capability that could potentially host the imagery. However, significant technological challenges related to network firewalls between the State Department and UNOSAT computer networks, hardware configuration, and web server performance eliminated that option.

In the few months that we were testing the UNOSAT option, progress on the HIU CyberGIS progressed to the point where it became the viable option for publishing the imagery. With the CyberGIS in mind, a modified workflow was built where HIU would vet incoming imagery requests using the criteria already established at the State Department to determine the viability of the requesting organization, purpose of the imagery, and whether having the imagery mapped into OpenStreetMap was useful. If yes, the HIU would obtain the NextView licensed imagery from an existing archive or submit a new request for tasking. Once received, the imagery was processed in-house, and at various points included geometric correction, subset and mosaic operations, contrast enhancements, and resolution merge operations. After processing, the imagery was transferred to the HIU CyberGIS platform and published as an image-based web service, specifically a Tiled Map Service (TMS) that could be directly read by OSM editing software. Once the TMS was published, it was integrated into the Humanitarian OpenStreetMap Team microtasking platform, called the OSM Tasking Manager. The reasons for using the TMS structure and for integrating with the OSM Tasking Manager are grounded in practical and legal distinctions. From the practical perspective, we needed to engage with the volunteer community using the tools and methods they were already using, and from the legal perspective, we were taking a very conservative approach and wanted to leverage the user name and password login structure implemented in the Tasking Manager.

Tiled Map Service

The use of a Tiled Map Service as the web publishing standard for IttC was based on three factors: it was an open standard that was already supported by OSM editing tools, it offered the best performance in terms of scalability and speed of delivering map tiles to the client application, and to meet the NextView and You criteria related to imagery-derived products. To understand the use of a TMS, a review of how other web services work is necessary. In a traditional Web Mapping Service (WMS) when the client application makes a request, *e.g.* a user zooms into a map embedded in a web page, the client application determines the bounding box, selected layers, projection, and other information about the map view and constructs a URL that contains all the information about the request. That URL is passed to the map server, which parses the URL to identify the relevant information, and begin the map rendering process. The map server extracts the needed data layers from the spatial database or file system and renders the requested data using the cartographic rules it is programmed with. The completed image is compressed into an image file (usually .png or .jpg), and sent back to the client for display. Note that at no point in this transaction does the server ever send the client the actual data, only the rendering of the data. This process is computationally intense as the map server and spatial database are used for every generated view. Benefits of this approach are that a user can change layers and styling dynamically, but has the downside of significant computational load and performance challenges for any system that scales to support many users.

The solution to this problem was to change the order of when map rendering takes place. Instead of waiting for a map request and then drawing the specific view, the idea was to pre-render every possible map view and store the pre-rendered map "tiles" for use later. With tiling, it is possible to pre-render a complex map with many layers then store or "cache" the tiles on disk. While this approach eliminates the user's ability to change layers and styles dynamically, it has significant performance advantages. By pre-caching tiles, the map server has to work hard up front to render and store all the tiles on disk, but at runtime, performance is greatly increased (1-2 orders of magnitude) as the computational overhead to deliver a map tile over the network is reduced to a file system call to identify and send the requested tiles (MetaCarta Labs 2010). This approach works well for complex basemaps where the data is fairly static and the goal is to have a consistent cartographic rendering, and for remote sensing imagery.

The development of tiled mapping schemas was a key element of GIS 2.0, and is the technology that the majority of modern web maps are based on. Google was the first to introduce the concept in Google Maps, and the world was immediately taken with the "slippy map" technology that allowed a user to pan seamlessly around the map. Slippy maps work because the browser is requesting tiles that are outside the current bounding box of the map and storing them in the browser's memory; this way if a user scrolls in any direction, the data is already on the client, and the process feels seamless. Additionally, map tiling is what made the "map mashup" possible, given a few caveats that are explained below.

The popularity of Google Maps meant it was becoming a de facto standard for web mapping. Beginning in 2006, the open source geographic software community recognized the need to be interoperable with Google Maps and began to standardize the parameters for building and consuming tiles. Out of the FOSS4G 2006 conference, the Web Map Service – Cache (WMS-C) specification was developed, which later morphed into the generic Tiled Map Specification (Kralidis 2008; Open Source Geospatial Foundation 2012). At the same time, support for TMS was added to the OpenLayers web mapping library, ensuring that web clients could consume the TMS format (MetaCarta Labs 2010; OpenLayers 2015).

It is this standardization that made the map mashup, the ability to combine data from different servers on the same web map, possible. It also allows different basemaps to be used across applications, an important enabler for the adoption of OpenStreetMap. There are several cartographic conventions that had to be standardized to make tiled map schemes interoperable. The baseline requirements involve a common reference grid for the entire world, known in cartography as a map projection, the tiles have to be of uniform size, and there has to be way of determining a common set of map scales that will be pre-rendered. These requirements are why the Web Mercator projection has become dominant in online maps, and why tiles are almost always rendered at the 256x256 pixels size. Beyond projection and tile size, the other requirement is to have a defined method for changing between scales, which in this case are referred to as zoom levels. While it is possible to create tiling schemes with any projection and tile size, because Google had a huge presence in the market, most everyone adopted their conventions and they became encoded in the baseline TMS standard.

In a TMS structure, the world is projected in Web Mercator (a rectangle with coordinates ranging from -85 to +85 and -180 to +180). At zoom level 0, the entire world is rendered as a single .png or .jpeg image with a dimension of 256x256 pixels, see Figure 26. Each increase in zoom level doubles in both dimensions, so at zoom level 1, the entire world is divided into 4 equal size tiles, each quadrant again rendered as a single .png or .jpeg image with a dimension of 256x256 pixels. As each zoom level increases, each previous tile is sub-divided into 4 equal size tiles, so zoom level 2 has 16 tiles, zoom level 3 has 64, etc. Using this structure, each location on earth exists in every zoom level, and the location can be described by a combination of zoom level, X, Y structure, see Figure 27 for an example. The finest level of resolution for mapping aerial imagery usually does not exceed Zoom Level 19, an approximate, latitude dependent map scale of 1:1,128. At zoom level 19, it would take 274 billion tiles to cover the globe; however, in reality that number is much smaller as ocean tiles are not cached. (MapBox n.d.; OpenStreetMap Wiki 2015d; Wikipedia 2015a; Open Source Geospatial Foundation 2012).

Figure 26 – Example of a TMS at zoom level 0 (MapBox n.d.)

Figure 27 – Example of a TMS at zoom level 1 (MapBox n.d.)

In order to satisfy the NextView requirements, the prototype IttC workflow was designed to be as conservative as possible, maintaining the strict approval processes that the HIU would use to share actual imagery, combined with the pre-rendering aspect of the TMS, and the user management controls of the OSM Tasking Manger (discussed below). Coming out of the RELIEF 11-4 meetings it was clear that in order for IttC to work, we needed to take particular care to protect the actual imagery pixels and to maintain the appropriate copyright. By using a Tiled Map Service we determined we could provide access to pre-rendered tiles of the source imagery that could be used for mapping, while simultaneously never providing direct access to the imagery data itself. As discussed above, the TMS only sends a compressed image file (or series of files) to the client for viewing. These individual map tiles, formatted as .jpeg or .png, cannot be reversed engineered to convert back to raw imagery. In every other use case in which satellite imagery is used to make a map that is then exported to a .jpeg or .png, the output product is considered an image-derived product, yet we continued to receive pushback that we were sharing imagery.

In order to perform the rendering for a TMS, a standard Web Map Service is involved. Essentially the tiling operation systematically walks through a WMS service, rendering each individual tile at the specific zoom levels. For the WMS to render, it does have to read the raw imagery, but after the tile is created, the raw imagery is no longer required. So in a final, almost absurd, level of protection we would only enable the WMS during the tiling operation, then shut if off while running the TMS, thereby absolutely ensuring that at no point was actual imagery every shared.

The final criteria we needed to uphold involved the expressly stated condition in the EULA that copyright be maintained with appropriate attribution. In a standard image-derived product, this could be a simple ©DigitalGlobe attribution in the corner of the map. For IttC we went a step farther in meeting this copyright requirement, and setup the TMS rendering to include the ©DigitalGlobe watermark on each tile in the cache, see Figure 28. By including the watermark as part of the rendering process, it is permanently embedded in each image tile, thereby exceeding the government's license requirement, and limiting to the maximum extent possible the risk that someone might try to download every tile and attempt to reassemble the data to sell for a commercial profit.

Figure 28 – Screenshot of IttC TMS service containing appropriate copyright markings on each map tile, note the faint ©DigitalGlobe marks on each tile. White arrows highlight three of these. Attribution watermarks are permanently embedded in each tile.

HOT Tasking Manager

The final step in the prototype workflow involved volunteer mappers accessing the TMS. To accomplish this, we leveraged the Humanitarian OpenStreetMap Team’s microtasking platform, the OSM Tasking Server (http://tasks.hotosm.org). The Tasking Manager is a web application that is used to manage volunteers working collaboratively on a specific map task, and is capable of managing hundreds of tasks at the same time. For each task, a description, instructions, tagging guidance, and organizations involved are included as well as a map interface for managing volunteers working on the task. As Figure 29 shows, the Tasking Manager sub-divides a large mapping task into a series of smaller tasks, providing a mechanism for a large number of volunteers to map concurrently without mapping over the top of each other. This is important as OSM does not have a mechanism to detect conflicts in real-time. The workflow starts with a volunteer selecting an unmapped tile, which will then automatically open the OSM data editor of their choice with the imagery TMS and spatial extent of their mapping area highlighted. Once mapping is completed, the user checks the data back into OSM and marks the task complete in the Tasking Manager, turning the square from grey to red.

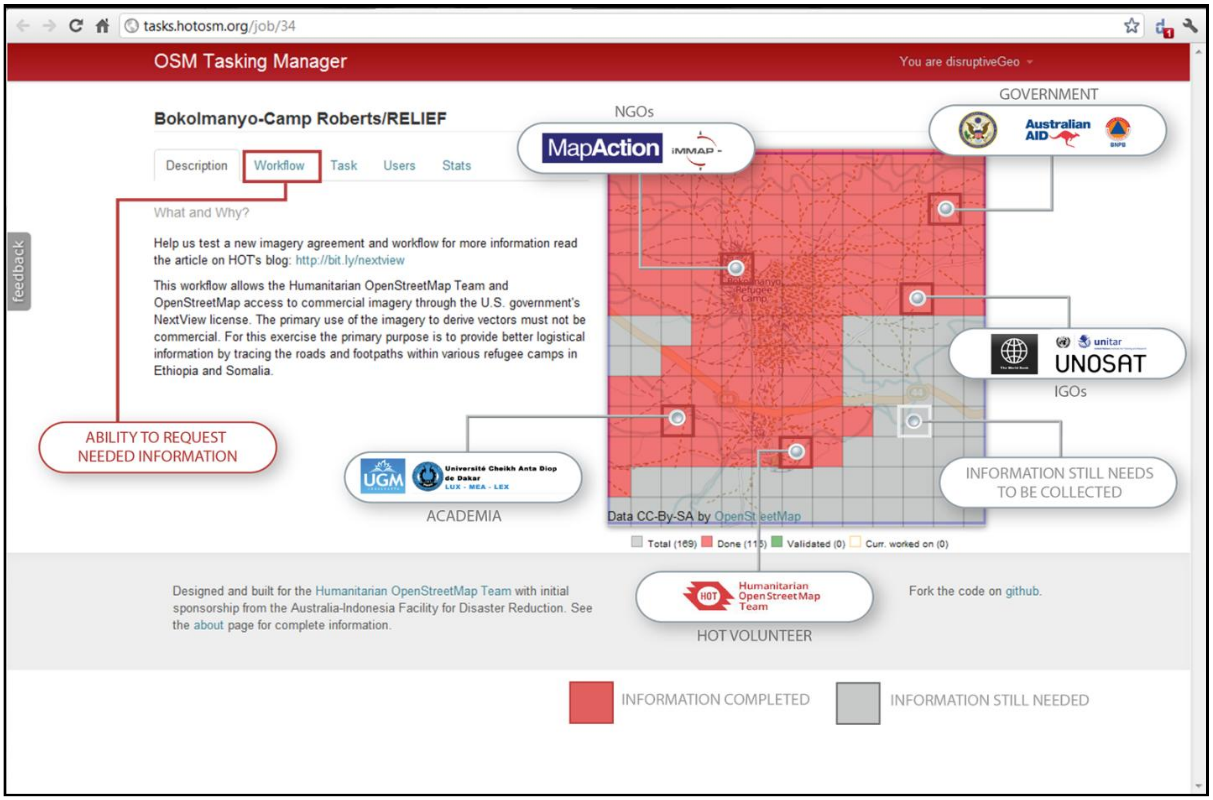

Figure 29 – Annotated view of HOT Tasking Manager Task page, depicts the map interface used to manage volunteer efforts and how individual sub-tasks appear as smaller grid cells. Grey indicates the tile still needs mapped, Red indicates its been completed.

Access to the Tasking Manager is managed using a volunteer's existing OSM user name / password login process, and if needed, the Tasking Manager can be configured to "white list" users to further refine who can access a task. Custom modifications were made to the Tasking Manager to accommodate the NextView license requirements, specifically, a mechanism to ensure that a volunteer mapper had to review and accept the NextView license acknowledgement before they could access the imagery for the task (Figures 30 and 31).

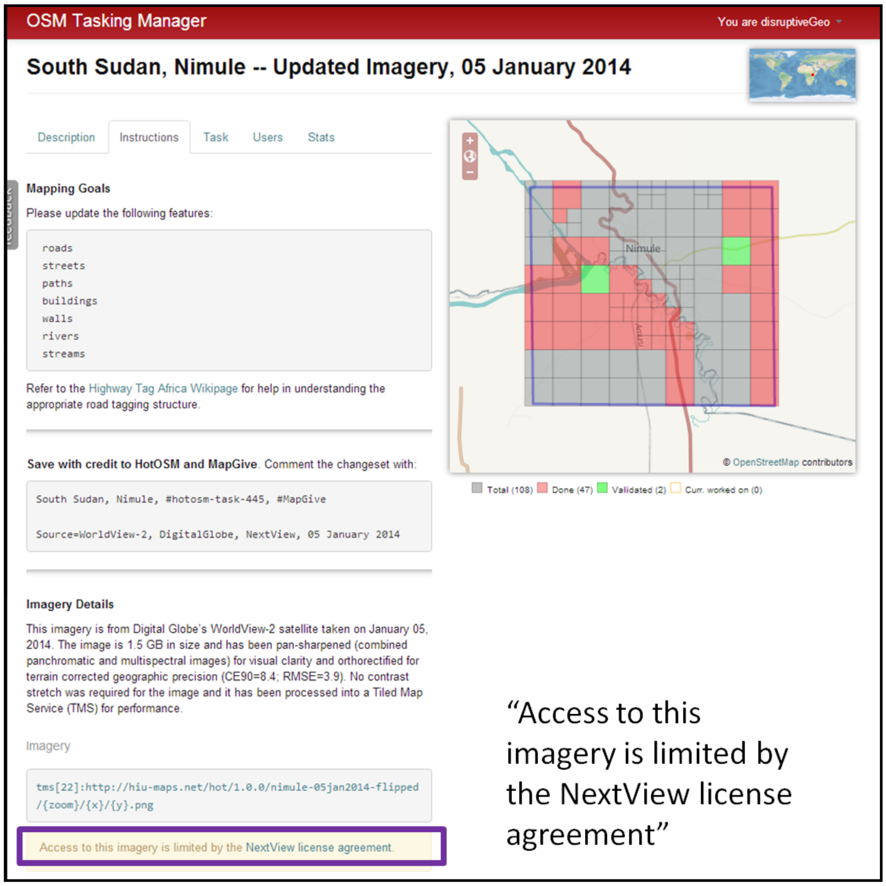

Figure 30 – Modification to OSM Tasking Manager to require a user acknowledge the NextView license before gaining access to the imagery.

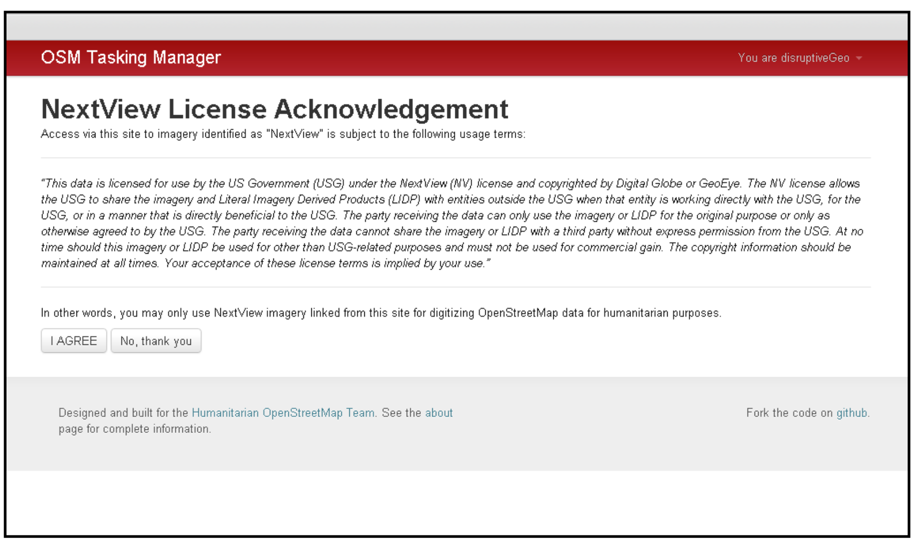

Figure 31: NextView License Acknowledgement integration into the OSM Tasking Manager platform

During RELIEF 12-2 held at Camp Roberts in February 2012, the prototype IttC workflow (vetted imagery requests, TMS enabled, watermarked, password protected behind the Tasking Manager, and NextView acknowledgement requirement) for sharing imagery over the Horn of Africa refugee camps with VTC for mapping into OSM was again presented to NGA General Counsel for review, and again, agreed to. The primary factor again was NGA's support of the Lead Federal Agency, and State Department's determination that the project was in support of USG interests. With this concurrence, the project moved forward, but progress was slowed by technological development at the HIU and challenges in obtaining appropriate imagery.

"NextView and You" and Proposed Definition Changes

In addition to the prototype workflow, a request to consider a change to the definitions used in the "NextView and You" document was also made to NGA General Counsel as part of RELIEF 12-2 in February 2012, and again a year later in January 2013. In my opinion, there are several revisions that are needed for the NextView and You document to address confusion regarding the Imagery to the Crowd methodology, and the specific criteria that define "Imagery" and "Image-Derived Product (IDP)". NGA General Counsel's conclusion that NextView only applies to pixels, and not IDPs is significant, yet the definitions used to describe these terms lacks the specificity needed when considering web service methods that emerged after the original document was released. The revised definitions provided below are intended to ground the distinction between Imagery and Image Derived Product in imagery science and the technical characteristics of sensor derived imagery. By using this perspective, the definitions allow for easier interpretation in the future as new formats and communication standards arise.

The Open Geospatial Consortium (OGC) has created a series of web communication standards for the transmission of geographic data over the Internet. This family of standards began with the Web Mapping Service (WMS) and has extended to include Web Coverage Service (WCS), and Web Tiled Map Service (WTMS). Each of these services requests, processes, and returns raster geographic data, including imagery, in fundamentally different ways. Based on the current definition of Imagery and IDP, there is not enough precision to account for the ways these services work. As an example, the WMS and WMTS, a variant of the TMS standard, both compress the requested image into a "dumb" format like .jpeg or .png before transmission back to the client and as such should be considered IDPs. Yet the WCS standard streams raw pixel values back to the client, allowing these streams to be used in additional analysis that leverages the sensor derived characteristics of the imagery, and as such, should be considered the same as raw imagery.

Referring back to the "NextView and You" document (Figure 24 and Appendix 4)

> **Imagery**
>
> Original
>
> Imagery is the image and associated metadata. Imagery can be further manipulated, enhanced, & processed. Example: GeoPDF, GeoTiff, NITF.
>
> Proposed
>
> Imagery: imagery in digital format with associated metadata. Imagery can be further manipulated, enhanced & processed, while retaining its original spectral and radiometric characteristics. Example: GeoTiff, NITF, or a compression thereof.
>
> **Image Derived Product (IDP)**
>
> Original
>
> Image Derived Product (IDP) – any product created from raw imagery – could include image metadata, but generally does not and often referred to as "dumb-image"

Proposed

Image Derived Product (IDP): a visualization made from digital imagery, but which has undergone some form of image transformation in which the original spectral and radiometric characteristics of the imagery cannot be reverse engineered. Can be in either an individual file format (JPEG, PNG, PDF or GeoPDF), or accessed via a web mapping service. Also referred to as a Literal Imagery Derived Product.

Non-Literal Imagery Derived Product

Original

Non-literal Imagery Derived Product is a product derived from the imagery but no longer looks like an image. Example: Line-drawings, maps.

Proposed

Non-literal Imagery Derived Product is a digital vector or product derived from the imagery. Example: digital vector files, line-drawings, maps.

IttC Prototype: The Horn of Africa Refugee Camps

Over the course of 2011 and 2012, famine raged across the Horn of Africa (ReliefWeb 2012). The occurrence of this famine corresponded with the technology development of the CyberGIS and the IttC initiative, and as such, made a natural test case. The previous CyberGIS section describes the construction of the "HoA Viewer", a geographic visualization platform designed to communicate the spatial and temporal patterns of the famine and its impact on human migration. The most dynamic element of the HoA Viewer was the daily updated refugee camp statistics that were harvested from United Nations High Commission for Refugees (UNHCR) documents. The visualizations and interfaces created to display these data in the HoA Viewer were impressive, but there was a problem in the geographic domain as the OpenStreetMap basemap we were using only had very limited information (see Figure 32). The commercial mapping companies actually had worse coverage with no data over the camps, and the Blue Marble satellite layer was too spatially coarse to see the camps.

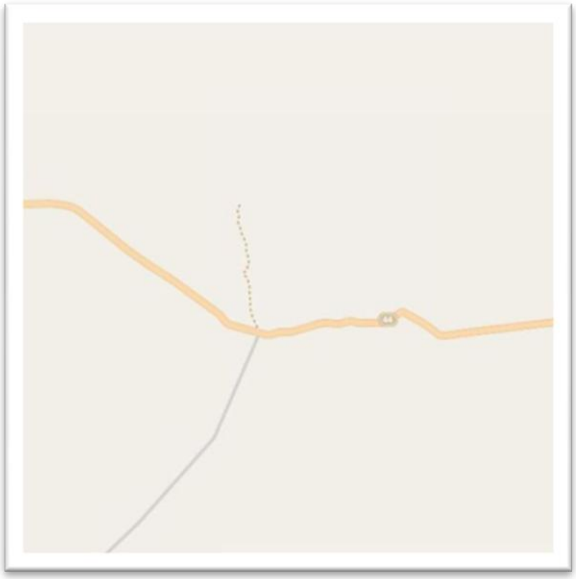

Figure 32: Screenshot of the Bokolmanyo refugee camp as depicted in the OpenStreetMap database before the HIU-led Horn of Africa Mapping Experiment (May 20, 2012) .

This presented a unique opportunity to test the Imagery to the Crowd methodology. On one hand it was a practical exercise. The lack of easily accessible, updated, GIS-ready data for the five refugee camps in Ethiopia, collectively known as Dollo Ado, and the five camps in Kenya, collectively known as Dadaab, meant that the HoA Viewer itself was limited in its utility as a visualization platform. The work would not be duplicative, for as far as we could tell, the data did not exist; as confirmed in personal communications with GIS staff at the United Nation High Commission on Refugees (UNHCR), the organization responsible for building and maintaining the camps. On another level, it was a moral issue; the inability to visualize the geographic footprint of approximately 600,000 people living in tents meant they were easy to forget. Their representation on the map was a critical element in understanding the situation.

By May 2012, the technological development of the CyberGIS had caught up with the legal and policy work on IttC, and it possible to use the HIU system in a prototype mapping experiment. As part of RELIEF 12-3, the HIU working in partnership with the Humanitarian OpenStreetMap Team held a mapping experiment. The goal was to produce detailed vector data for the refugee camps, including roads and footpaths in and around the camps. With very limited advertising, consisting of one short blog post and two tweets, the goal was to determine if the volunteer mapping community would respond to a humanitarian mapping request from the HIU. The answer was yes, and over the course of three days, a large portion of the ten refugee camps were mapped. The total amount of mapping was impressive, as illustrated in Figures 33 and 34, as was the fact that less than 35 volunteers participated.

The results of the experiment yielded two positive outcomes, first it provided the evidence we needed to convince internal policy makers that this process did work, that the crowd was capable and willing to contribute their mapping efforts, and that the resulting data was of sufficient quality to meet the needs of multiple stakeholders. Second, the integration of the high resolution imagery services and the updated OSM data over the refugee camps significantly increased the utility and presentation value of the HoA Viewer, providing a vote of confidence to the CyberGIS initiative.

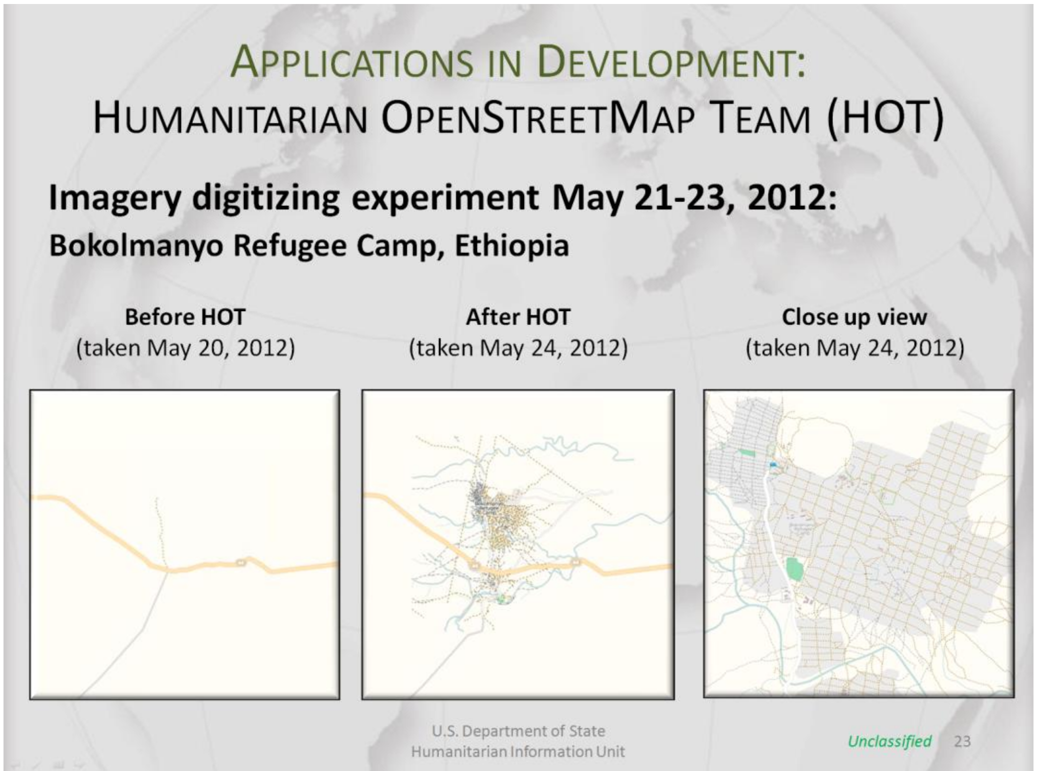

Figure 33 –Sequence of images showing the OSM basemap over the Bokolmanyo refugee camp during the IttC prototype

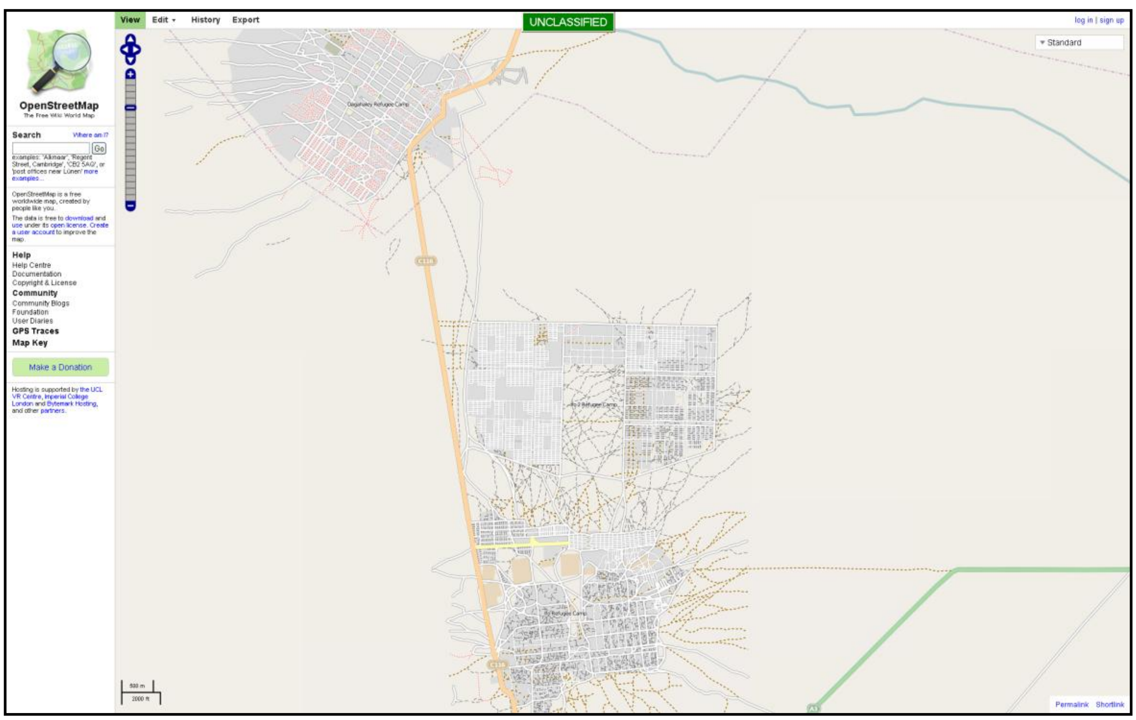

Figure 34 – OSM basemap over the Dadaab refugee camps (Hadadera, Ifo 1 and 2, and Dagahaley) after the IttC prototype

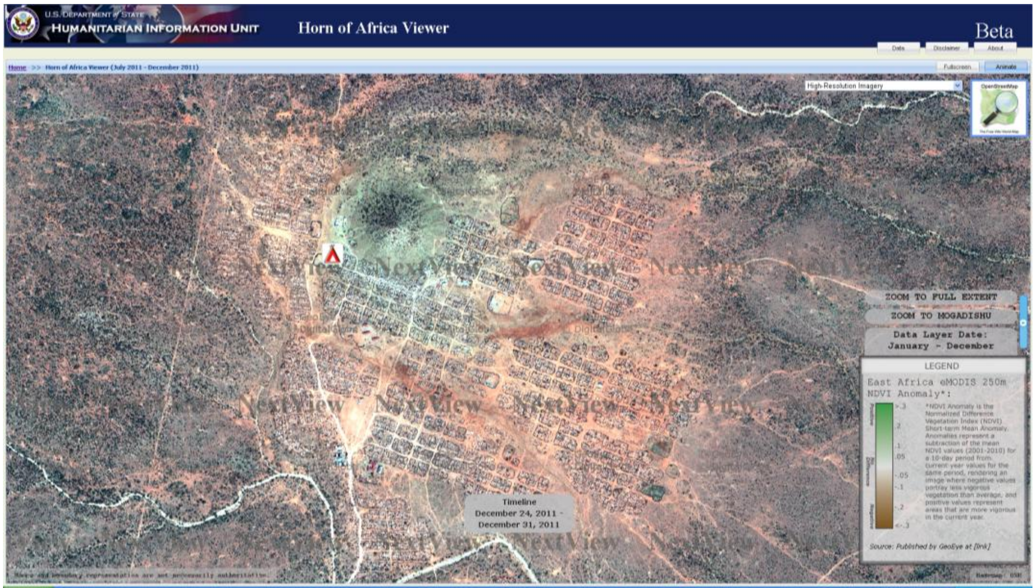

Figure 35 – Incorporation of NextView imagery services used in the IttC prototype into the HoA Viewer

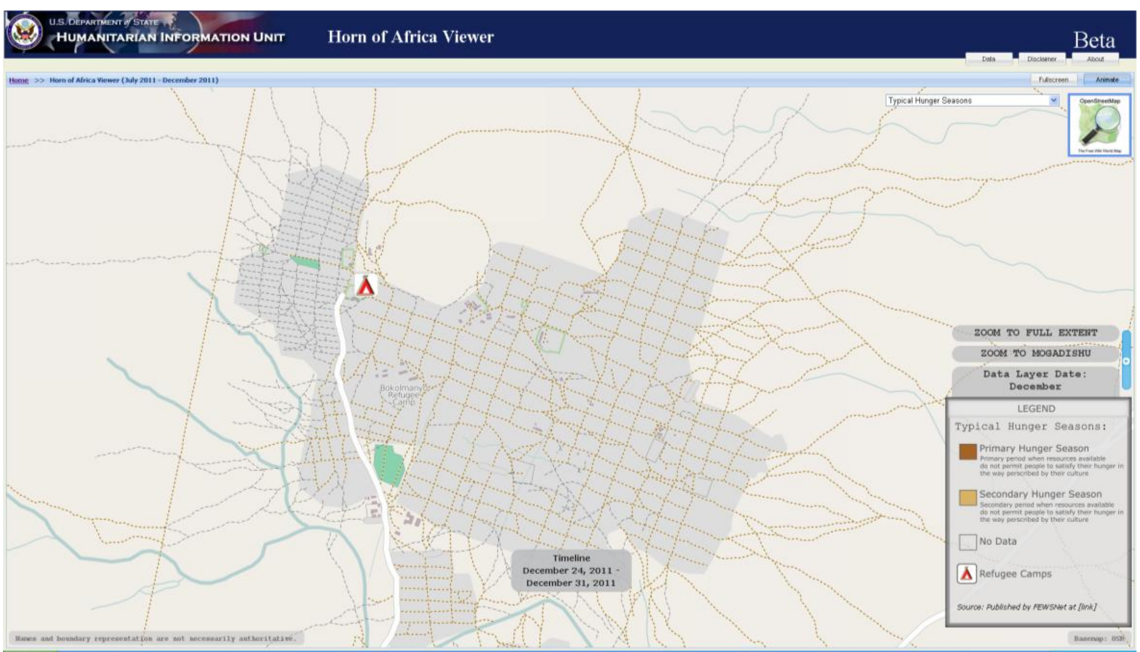

Figure 36 – Results of IttC prototype visualized in the HoA Viewer

Results

Results of four other IttC projects are also presented here. They occurred in different areas of the world, for different humanitarian purposes, and engaged different partners. Collectively they represent a good cross-section of the utility of OSM data and the value of IttC to support it.

Community Resilience – Gulu and Lira, Uganda – American Red Cross

This project was the second IttC deployment and supported an American Red Cross project focused on community resilience in two cities, Gulu and Lira, in northwest Uganda. Both cities had grown significantly in the last couple of decades and there were no good road or fire hazard maps available. The Red Cross project involved training local university students in the cities on how to map in OSM and how to conduct on the ground surveys to validate their mapping effort. A remote mapping party was held in DC to help the project by mapping as much imagery as possible, so that the field teams could focus on validation.

The biggest risk in both cities were the large numbers of circular thatch huts built in close proximity to each other and other residential structures. Figure 37 below shows what the raw OSM data looked like, and Figure 38 shows the rendered basemap, as it appears on the OSM website. In both examples the large numbers of huts are easy to detect. As part of the field survey portion, the local teams also collected business names and other details not visible in the imagery. In one of my favorite examples, there is a restaurant mapped in Figure 38 called Alulululu's Pork Roasting Joint. Note both of these cities were essentially blank slates in OSM before this project started.

However, it is Figure 39 that speaks to the potential of these projects to improve people's lives. This map is a simple extraction of the OSM road network in Gulu, mapped in grey, with all the huts that were mapped displayed in red. This map of fire risk is an extremely valuable tool for the local first responders, and since they were also trained by the Red Cross on OSM mapping, they have the tools and know-how to keep this data about their community up to date.

Figure 37 – OSM data displayed in the JOSM editor, note the large number of small circular huts.

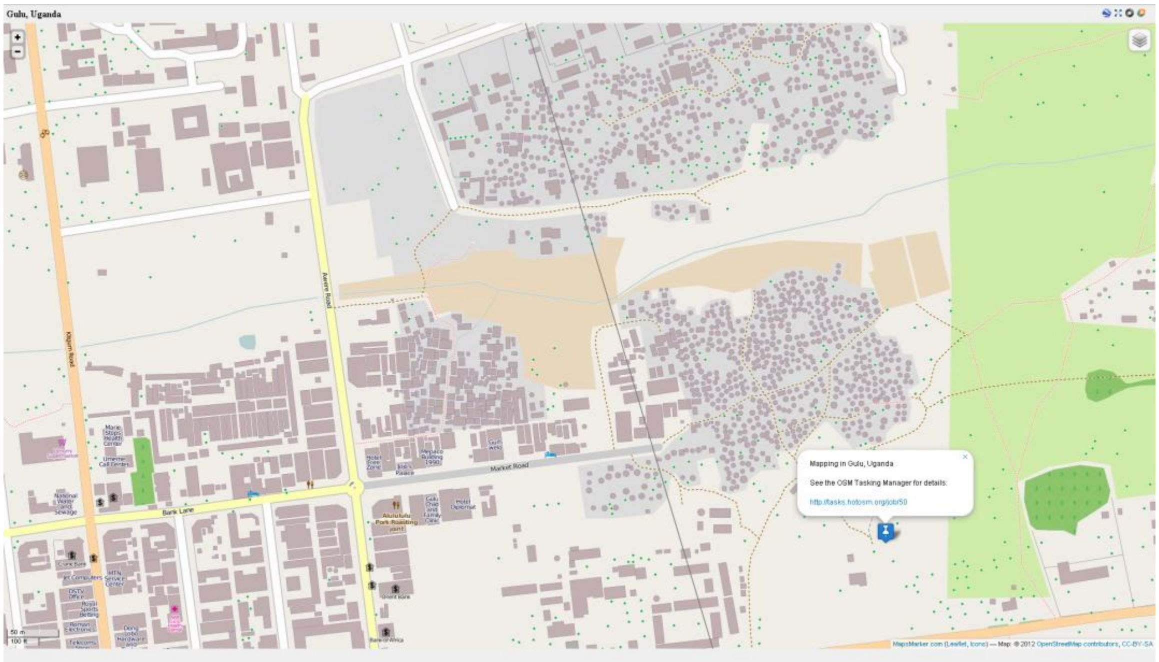

Figure 38 – OSM rendering of the data, displaying number of huts, trees, and business names.

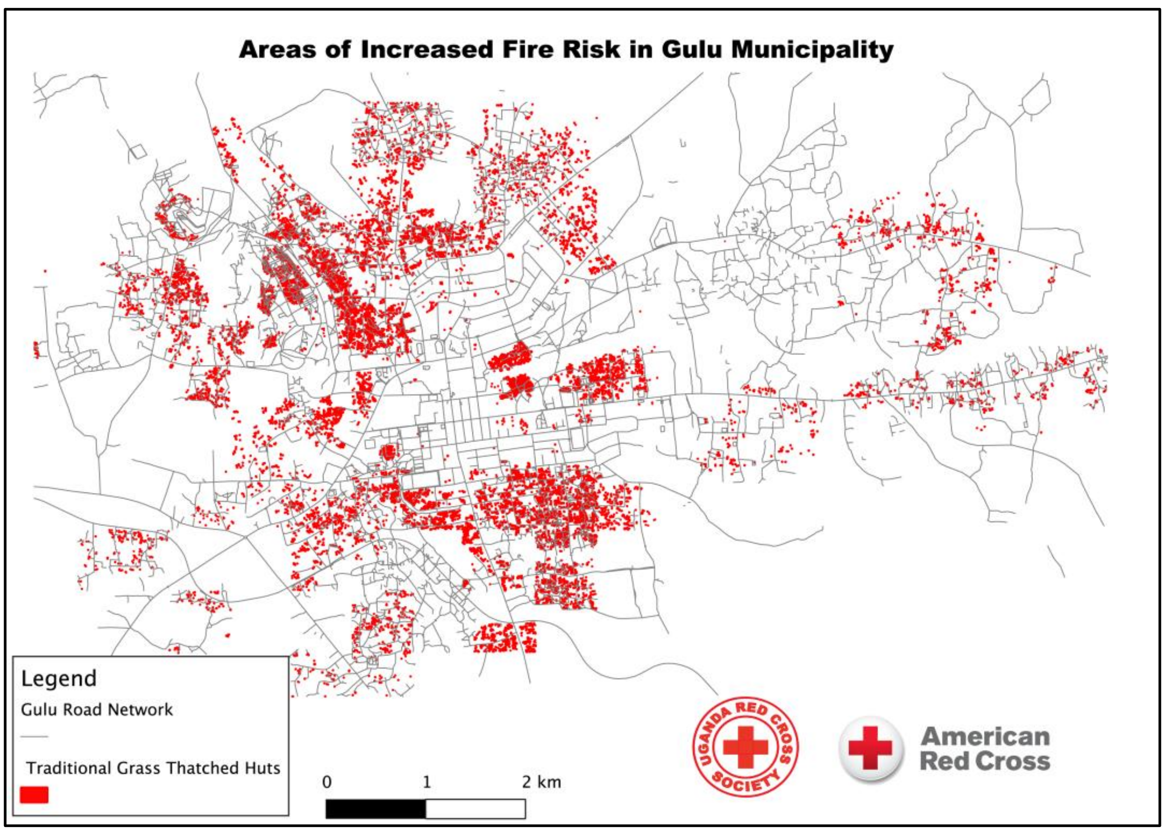

Figure 39 – OSM based map of fire risk in Gulu, Uganda

Disaster Risk Reduction – Kathmandu, Nepal – World Bank Global Facility for Disaster Reduction and Recovery

This project in Kathmandu originally started in February 2013 and ran continuously for the year. Conducted in support of a World Bank GFDRR project entitled OpenCities (World Bank and Humanitarian OpenStreetMap Team 2014), the goal of this project was to generate the base geographic data needed for seismic risk analysis in the Kathmandu Valley. As a two part process, IttC would provide an updated imagery service for use with remote mapping events, or mapathons, with the intent to map all the building footprints in Kathmandu. In the second step, a local Nepalese NGO called Kathmandu Living Labs would conduct ground surveys and validation of the OSM data for critical infrastructure and health facilities, annotating the OSM building polygons with additional information about building age, height, materials, etc. The annotated data would be used in seismic risk models. The USAID GeoCenter was also supporting the project and helped conduct one of the first IttC mapathons in conjunction with George Washington University in November 2013 (Roberts 2014).

Figure 40 – Screenshot of the Open Cities website, note dense urban environment

The density of Kathmandu makes for a challenging mapping environment, and this task tested the limits of imagery quality, volunteer skills, and CyberGIS capacity. Figure 41 shows the rendered OSM basemap for the central district of Kathmandu.

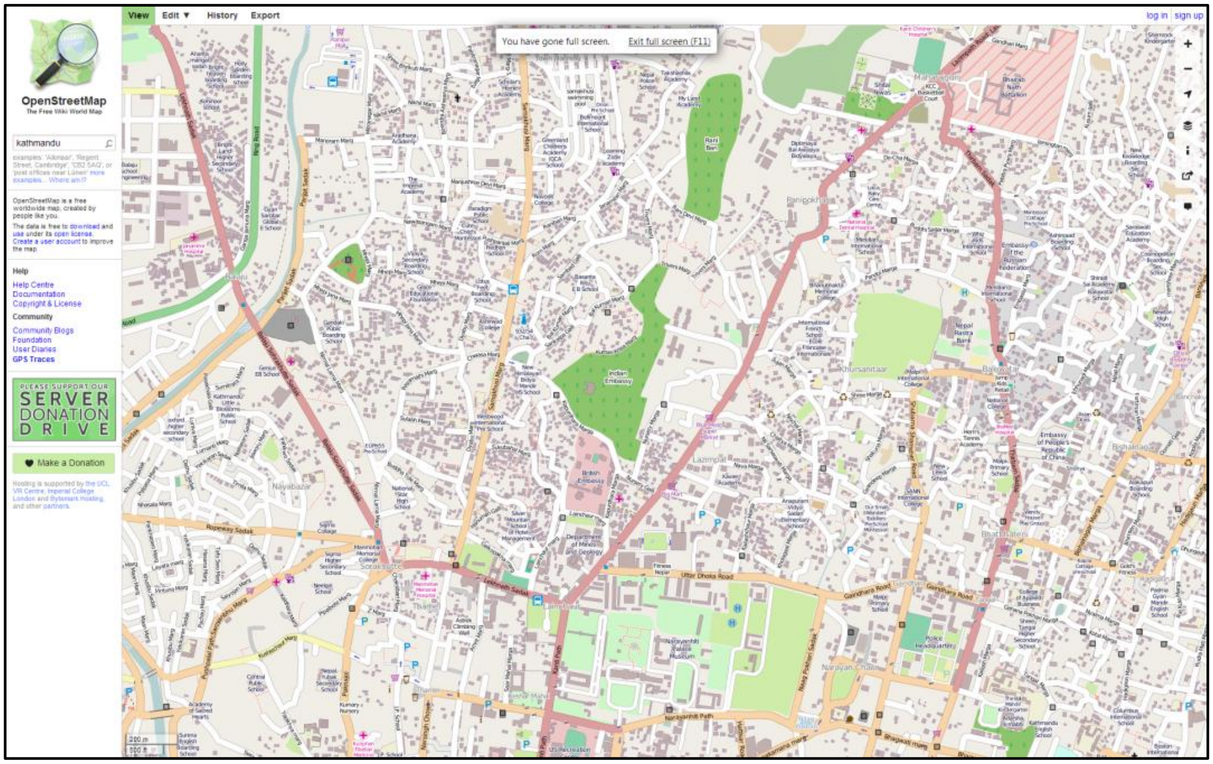

Figure 41 – OSM rendered basemap for central Kathamandu, showing hundreds of building footprints

In total, this project mapped over 118, 000 buildings in Kathmandu and the surrounding area, with 6,000 schools and health facilities ground validated and annotated with additional building information. Figure 42 displays the mapped buildings, color coded with the date they were mapped. The areas of light blue and orange / red represent large mapathons that focused attention on the task. From the IttC perspective, this task taught us a lot about how to scale the CyberGIS to support a large number of concurrent mappers, and introduced us to the mapathon concept. This experience would help shape the education and outreach elements created as part of MapGive.

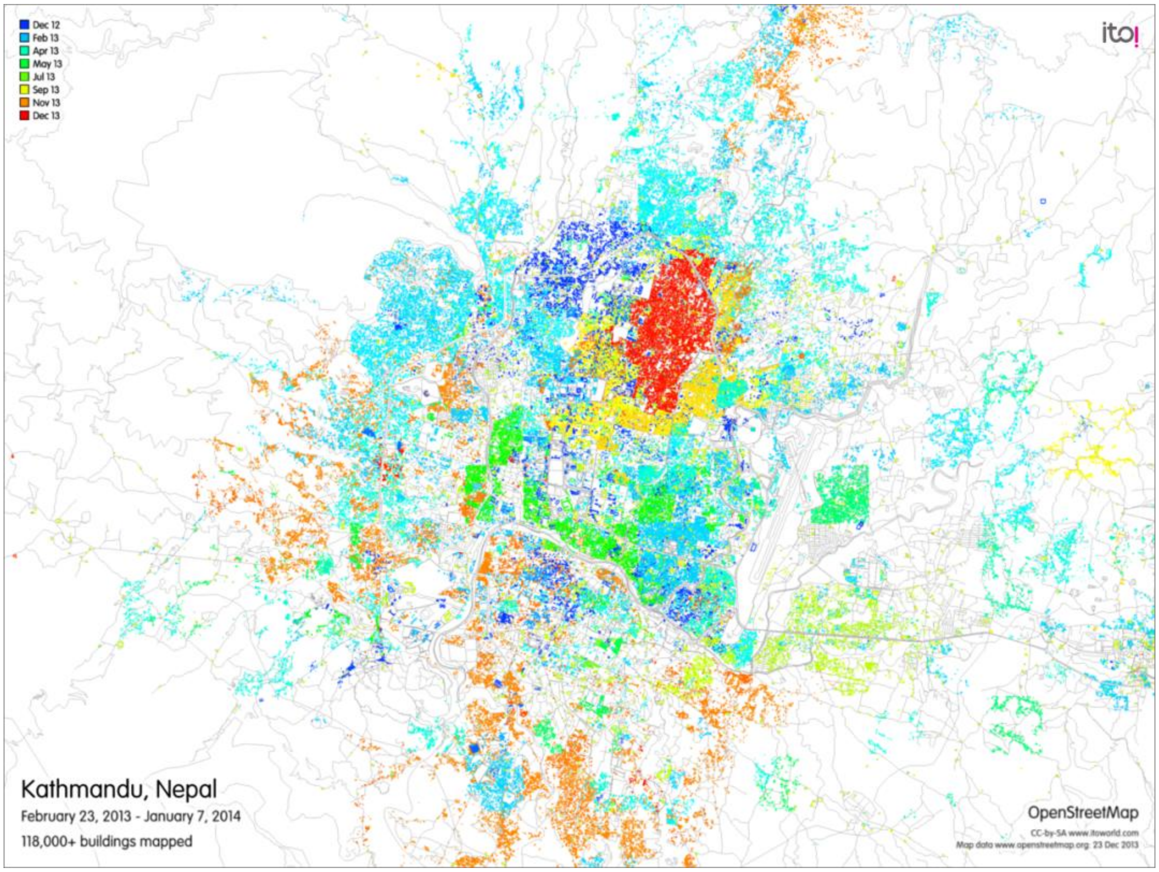

Figure 42 – Rendered building footprints in Kathmandu, Nepal, color coded by date of mapping

Kathmandu also marks the first time that a true disaster struck in an area that had previosly been mapped as part of a disaster risk reduction project. Clearly the Nepal earthquake is a disaster, but for once it was nice to already have the data in hand, and gratifying to know that the Kathmandu Living Labs team is on the ground coordinating response efforts using this data and not spending time mapping the city. As discussed in the Introdution, the fact that Kathmandu was already so well mapped meant that MapGive could focus its efforts in another affected area.

Disaster Response – Typhoon Haiyan in the Philippines – American Red Cross

Super typhoon Haiyan / Yolanda made landfall in the Philippines with wind speeds in excess of 170 miles and hour, and is the deadliest Philippine typhoon on record with over 6,000 killed. The response to Haiyan had several unique characteristics including the HOT pre-activation, the coordination of imagery inside the USG, the speed of the IttC deployment, and the activation of a unique clause in the EnhancedView contract.

For starters, the typhoon made landfall on Thursday evening eastern time on November 7, 2013, which was the day before the Veteran's Day holiday, which meant DC was empty and the government was closed until Tuesday of the next week. It was also my wife's birthday, and we were on a weekend vacation in Delaware. However, everyone knew the typhoon was going to be massive and started preparing. HOT had taken the unique step of formally activating a couple days before Haiyan made landfall; the projected path (Figure 43) was fairly confirmed and they started producing very detailed pre-event data along the projected path using Bing imagery. In reviewing Bing there were some gaps in data, so I started pulling down pre-event imagery in case we needed to fill those gaps.

In the three days after landfall, cloud cover precluded any commercial satellite imagery acquisition. However, the disaster response group at NGA had activated and was using national technical means to evaluate the storm damage. They had already received permission to release the damage estimates and pushed out the first set of data to humanitarian organizations on Saturday afternoon, Nov 9 (Figure 44). At some point on Sunday, Nov 10, the first commercial satellite images were shot, becoming available that evening. NGA released another set of damage assessments in the early morning of Monday, Nov 11. Additionally, the Copernicus emergency management group in the EU had also received the imagery and published their damage assessment early on Monday, Nov 11 (Figure 45). Given that the processing pipeline for ortho-rectifying imagery and putting it into the online archive takes about a day, by 6pm on Nov 11, I began downloading the post event imagery over Tacloban. Heading back into the office on Tuesday, Nov 12, we worked all day to process the post-event imagery over Tacloban, and pre-event imagery over Ormoc City, releasing both tasks to the OSM Tasking Manager at 1am Wednesday Nov 13. Within 36 hours the HOT community had completed mapping on both tasks, having mapped 21,000 damage and destroyed buildings in Tacloban (Figures 46, 47, and 48), and a complete pre-event map of Ormoc City (Figures 49 and 50). Two days later on Friday Nov 15, we published a third service over Cebu Island, and a couple weeks later published a fourth task for Carles.

On Saturday, Nov 16 the NGA, at the request of USAID, USGS, and State Department, executed the new "H.22 (U) Emergencies, Disasters, And Humanitarian Efforts" clause in the EnhancedView contract (Appendix 2) that allowed for a blanket release of all NextView licensed imagery to anyone for 30 days. To my knowledge this was only the second time the H.22 clause had been approved, and over the next 30 days the U.S. Government provided free and open access to over 1,500 high resolution satellite images covering the entire affected area through the USGS online portal, HDDS.

By the end of the Typhoon Haiyan response, almost 1,700 volunteers made 4.8 million map edits to the OSM database, which at that point was by far the largest volunteer response (Figure 51). For comparison, 600 volunteers mapped for Haiti and 4,700 and counting for the Nepal earthquake.

In my opinion, the Haiyan response demonstrated how the response system should work with regards to imagery and response coordination. Recapping the steps, first HOT activated before landfall, creating pre-event base data using Bing. When cloud cover precluded commercial imagery after landfall, NGA provided the first set of damage assessments using national systems, and publically released the data. As commercial became available, a professional analysis from the Copernicus research group was published at a finer spatial resolution than NGA's, and the IttC services provided both pre-event imagery, to fill a gap in the Bing coverage, and post-event imagery in the hardest hit areas. Finally, as the magnitude of the disaster began to overwhelm the collective capabilities, the USG opened up the entire collection of commercial imagery to be used for free by anyone. All of this occurred with the first seven days after the response.

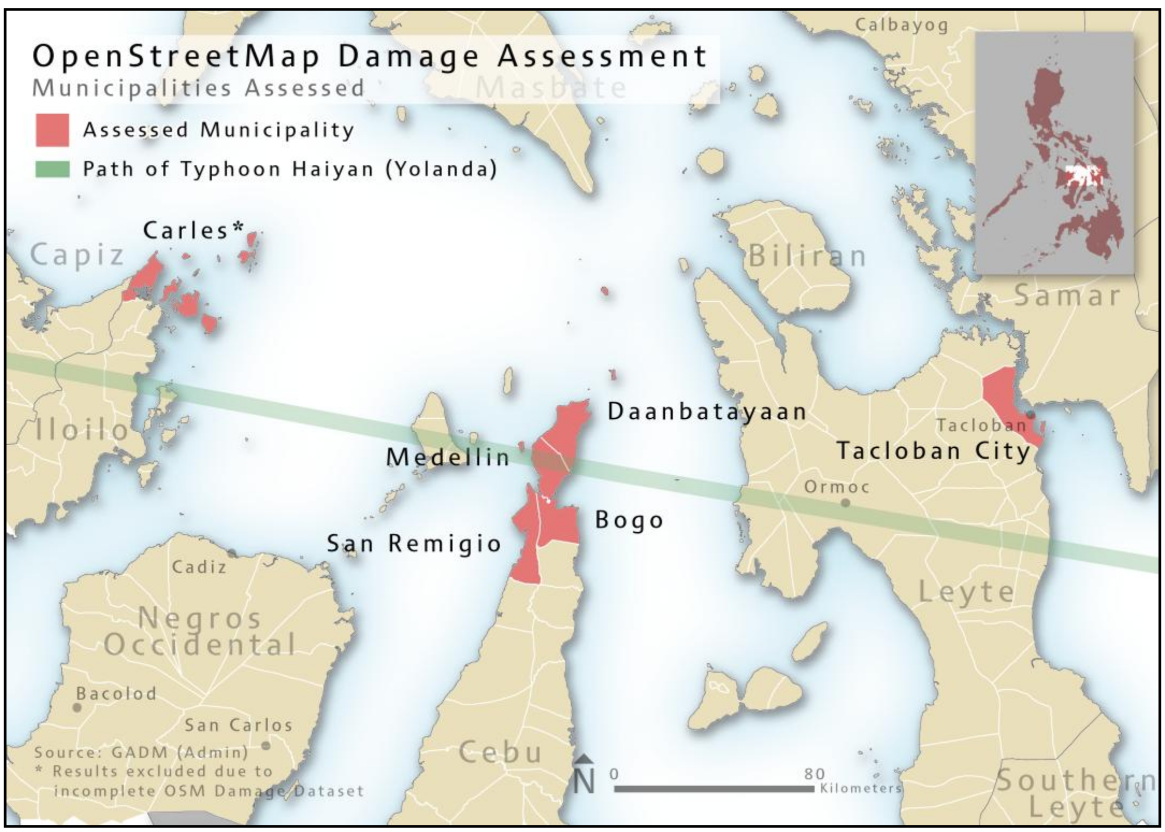

Figure 43 – Map of Typoon Haiyan path, courtesy American Red Cross

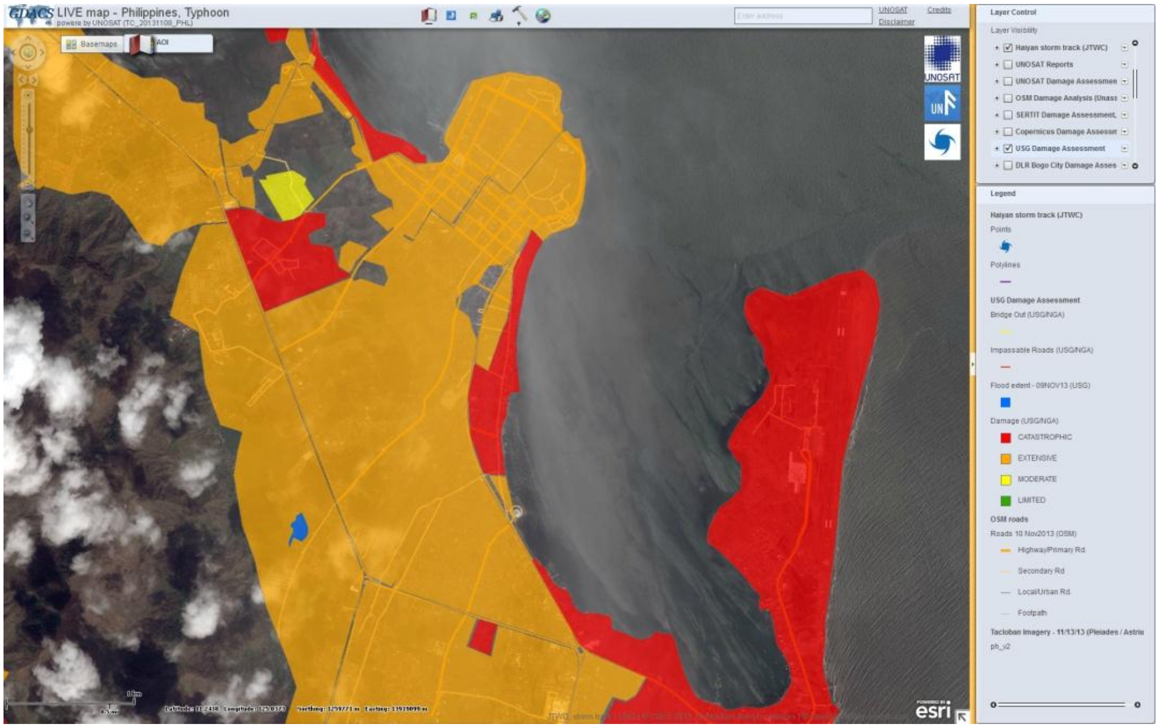

Figure 44 – Map of NGA damage assessment for Tacloban (Nov 9)

Figure 45 – Map of Copernicus damage assessment (Nov 11)

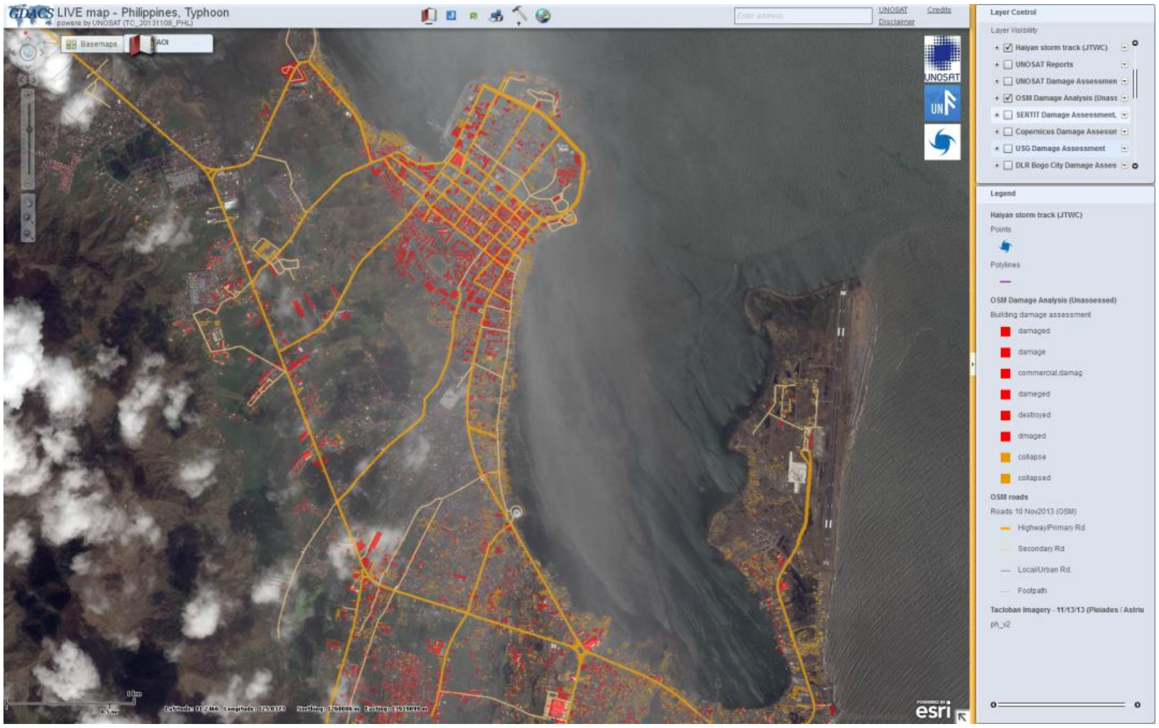

Figure 46 – Map of OSM damage assessment (Nov 15)

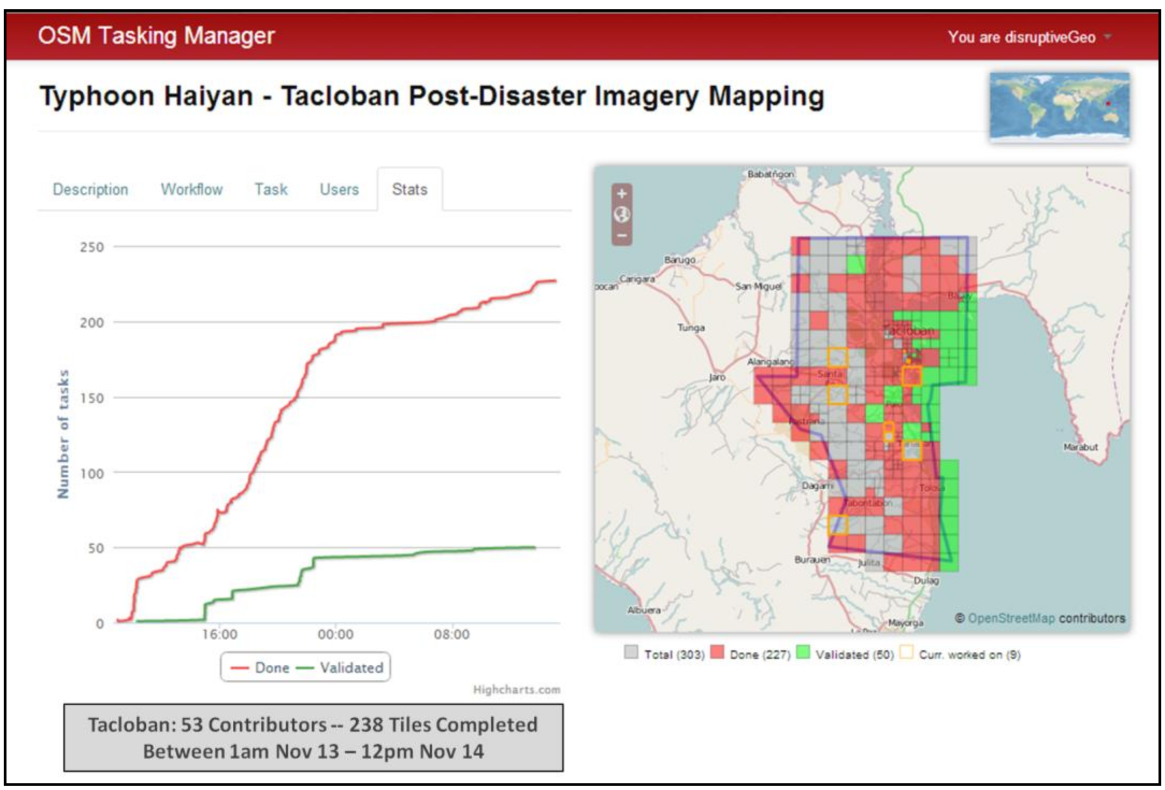

Figure 47 – Coverage and rate of completion of Tacloban task

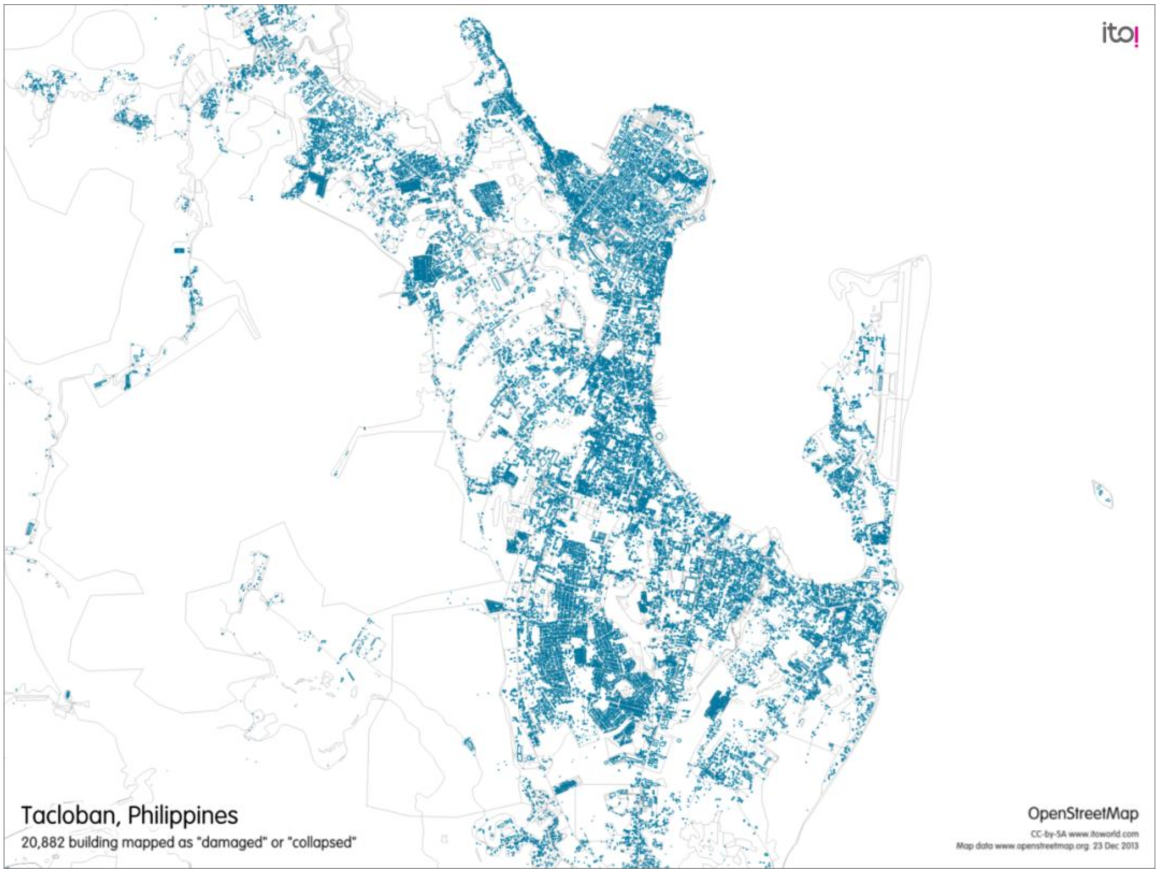

Figure 48 – OSM damage assessment, buildings tagged as damaged or destroyed

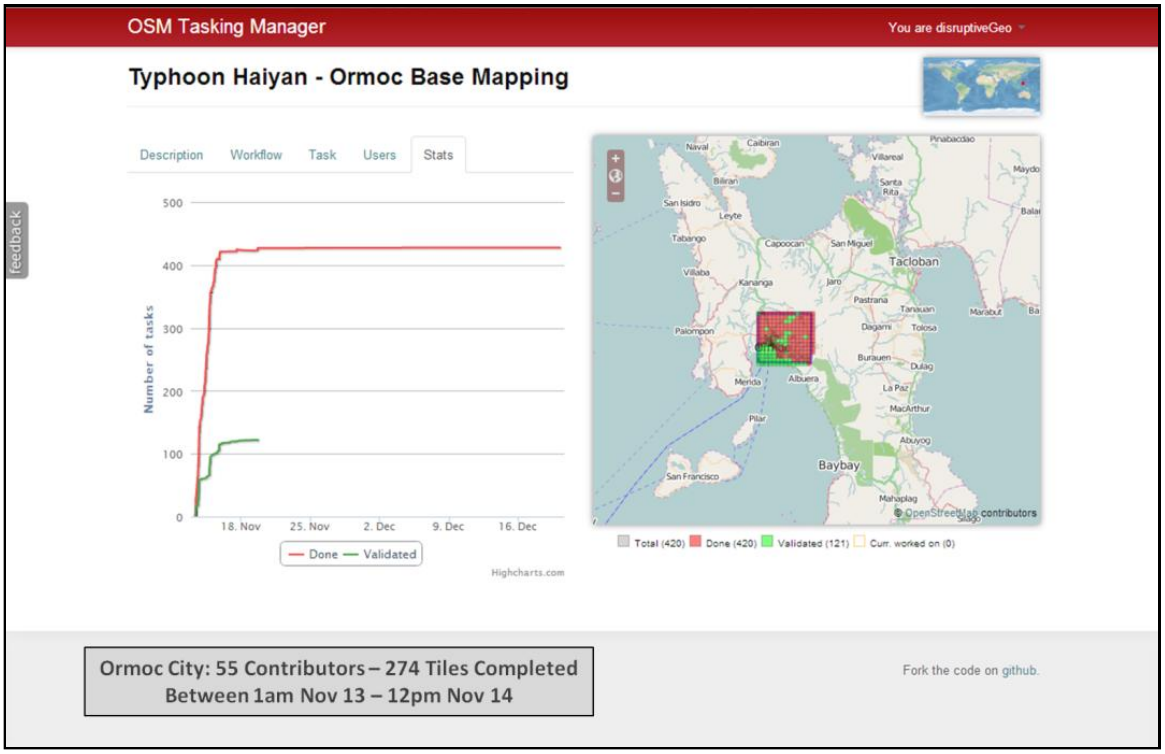

Figure 49 -- Coverage and rate of completion of Ormoc task

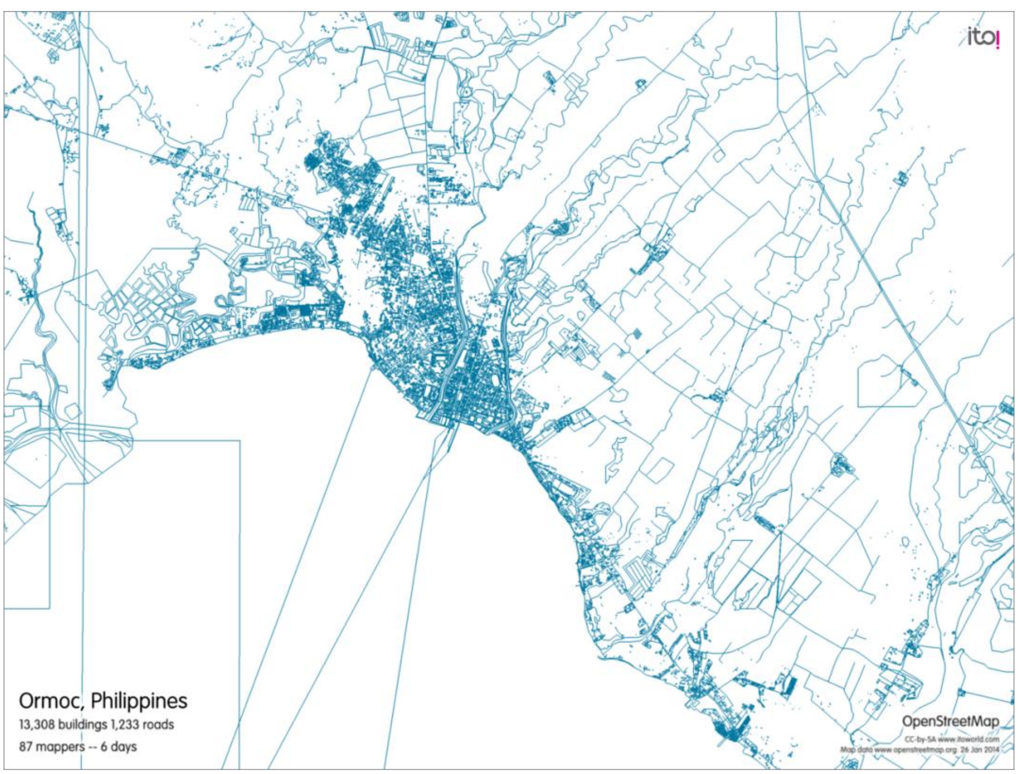

Figure 50 – Outline of Ormac pre-event data

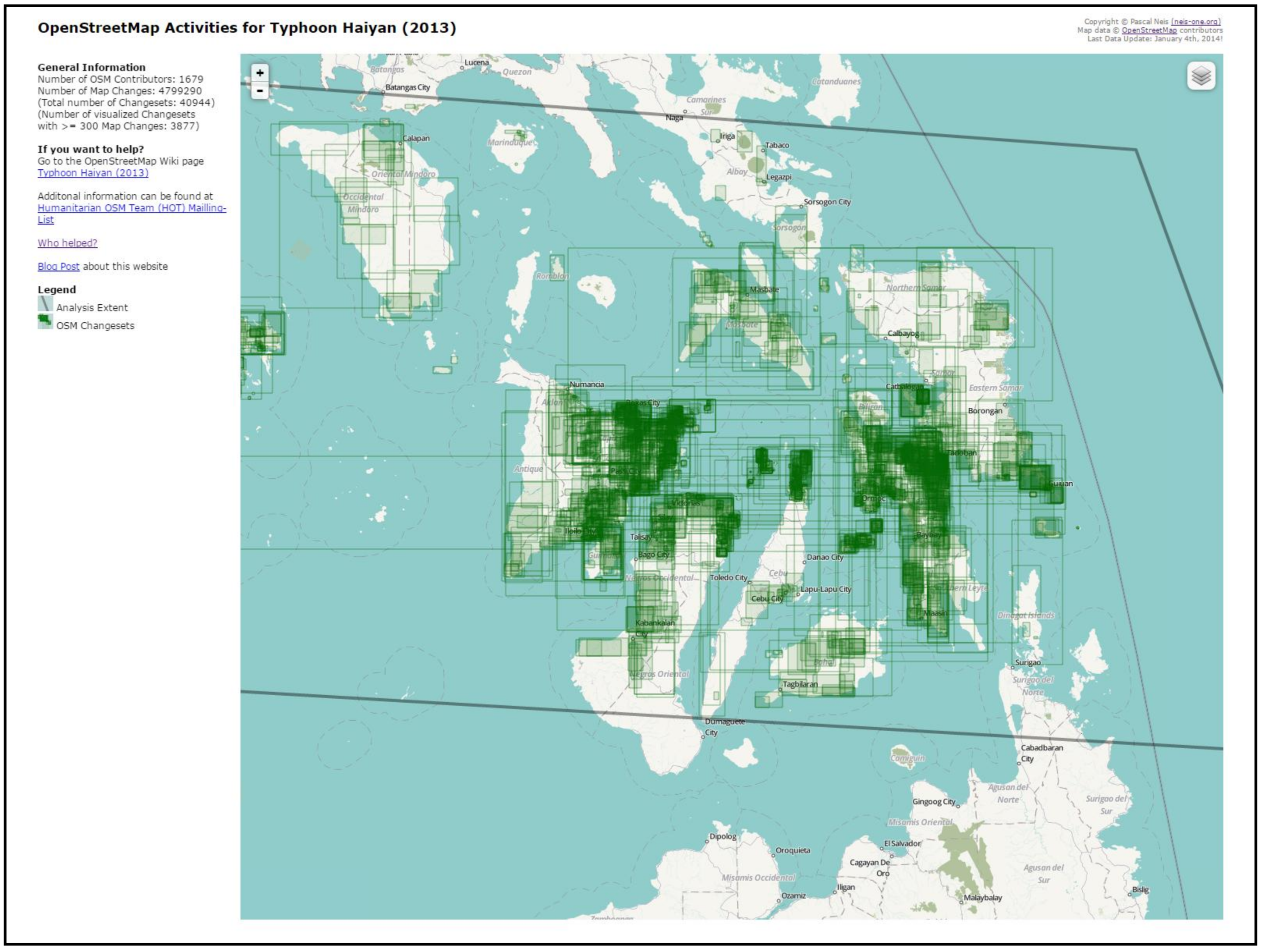

Figure 51 – Image showing all the map edits and statistics for the Typhoon Haiyan response. In total, 1,679 mappers made 4.8 million map edits, the largest activation up to that point.

Humanitarian Planning – Kindu, Democratic Republic of Congo – UN OCHA

This project was in support of humanitarian planning activities for UN OCHA, as they intended to increase attention in this region of DRC. The remarkable thing about Kindu is that is speaks to the amount of potential volunteer mappers. The release of this task was scheduled, there was no emergency, but that timing coincided directly with the Typhoon Haiyan response in November and then the eruption of violence in South Sudan in December, both of which attracted significant volunteer attention. As the graph below shows, the pace was constant through it all. This was also one of the first events where we thought to actually take good before and after images, which are impressive. Kindu was basically two roads at the beginning, and ended with 2,800 roads and 43,000 buildings.

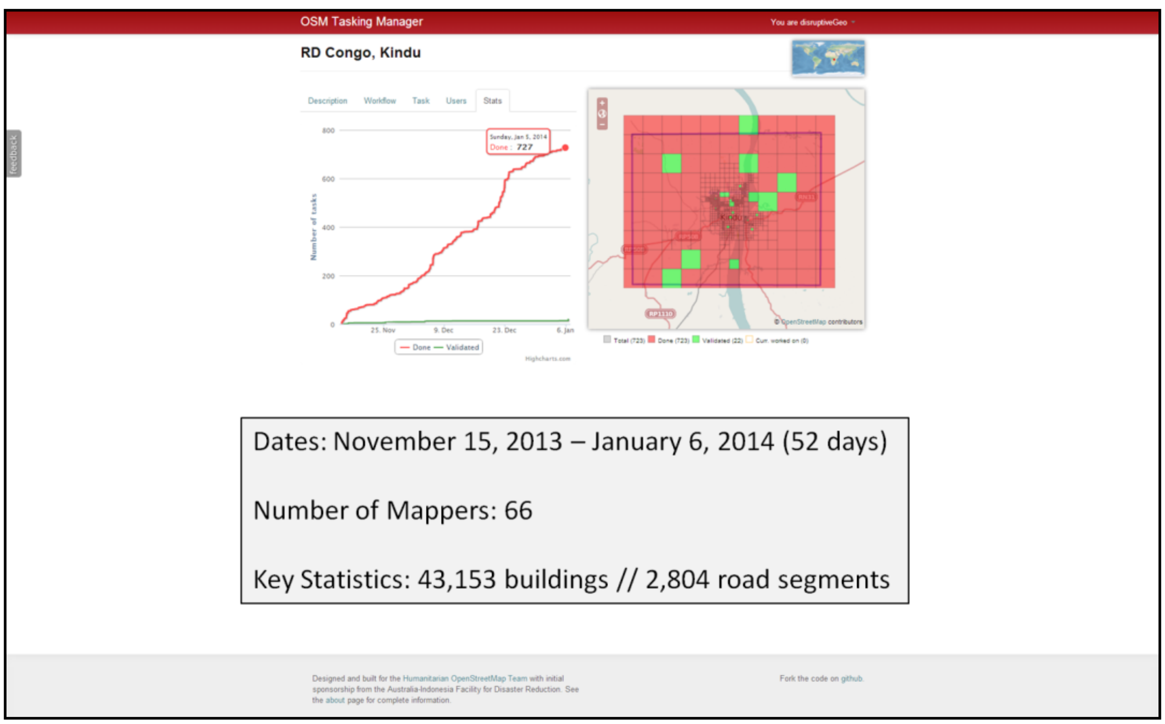

Figure 52—Screenshot from OSM Tasking Manager and statistics for completion of Kindu task

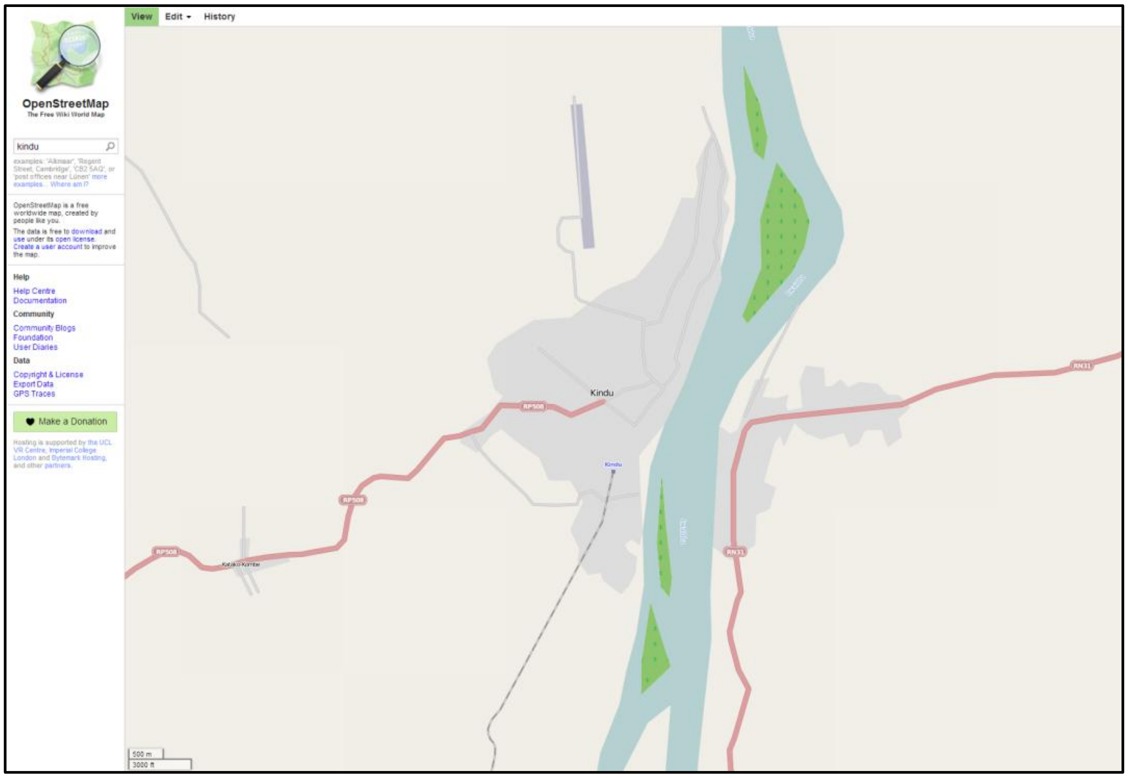

Figure 53 – Kindu, before mapping, November 8, 2013

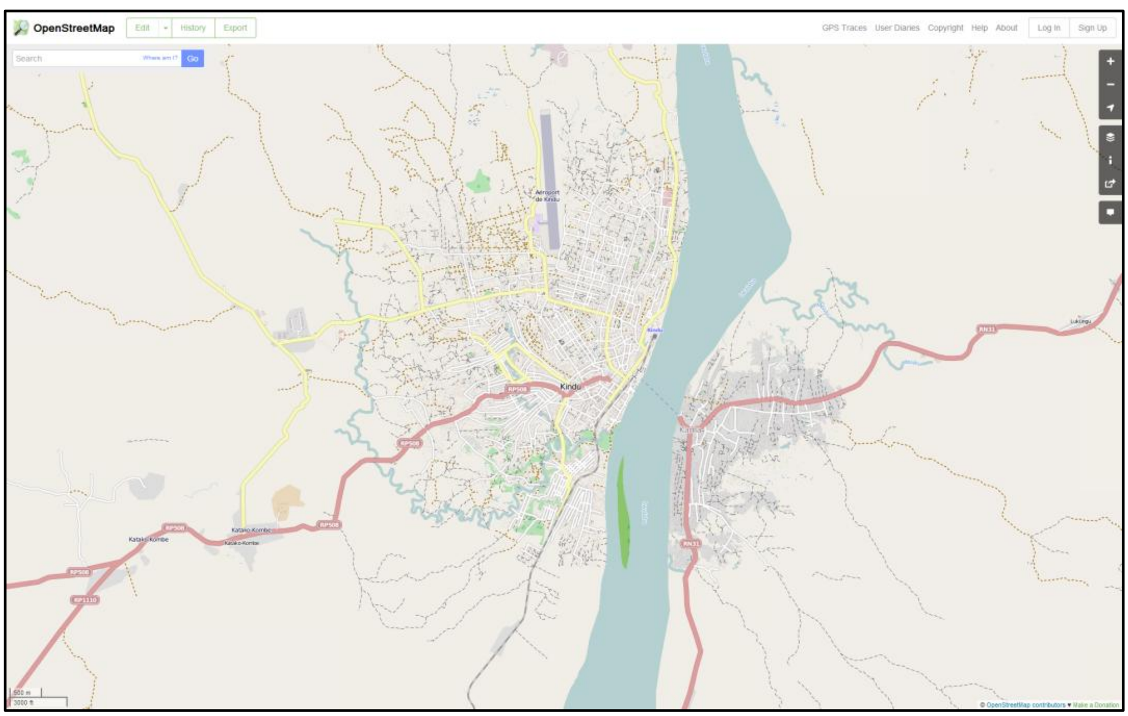

Figure 54 – Kindu, after mapping, January 6, 2014

http://www.openstreetmap.org/#map=14/-2.9480/25.9175&layers=N

Figure 55 – Kindu city center, before mapping, November 8, 2013

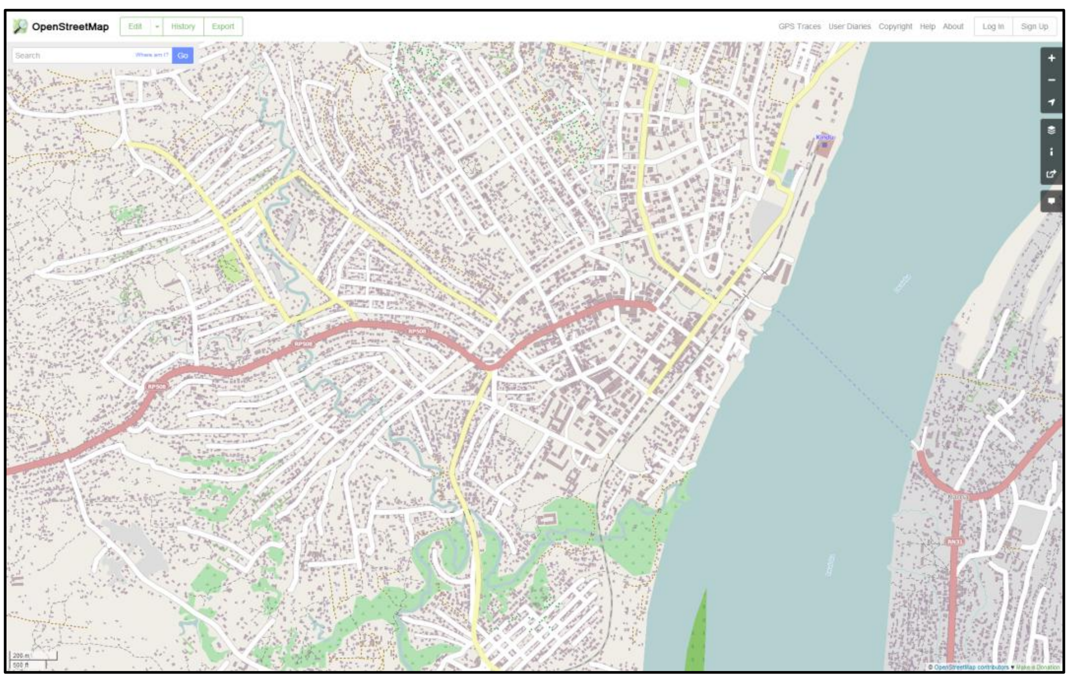

Figure 56 – Kindu city center, after mapping, January 6, 2014

http://www.openstreetmap.org/#map=16/-2.9503/25.9196&layers=N

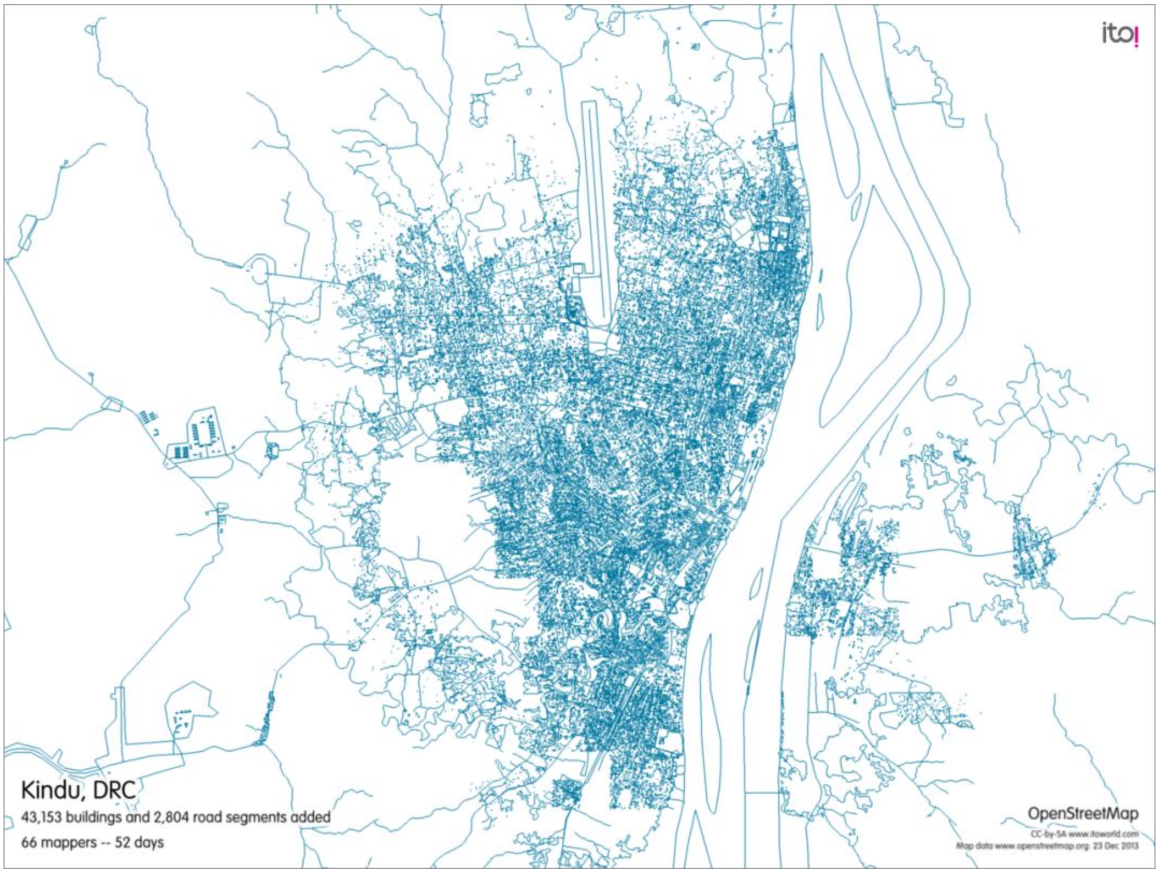

Figure 57 – Outlines of all features created for Kindu task

IttC Interactive Map and Data

An updated, interactive map of the areas where IttC has been deployed is available on the MapGive website (http://mapgive.state.gov/ittc/), along with a downloadable version of the imagery extent dataset. An analysis of the footprint dataset indicates that approximately 62,000 square miles of imagery has been made available to the HOT community through IttC. It is unlikely that we can calculate a cost to USG for the imagery (as the internal pricing per square mile for EnhancedView is redacted on the official DigitalGlobe SEC filings), but based on DigitalGlobe's published rates (CartoDB 2015), the total value would be in the range of $1.5-2M.

Figure 58 – Interactive map display the IttC imagery footprints

MapGive

In collaboration with the Office of Innovative Engagement (OIE) in the Bureau of International Information Programs (IIP), the HIU set about building an educational and outreach layer around the IttC core. The need for MapGive had become apparent after IttC participated in the Kathmandu mapathons. These events demonstrated there were a lot of people interested in humanitarian mapping, but that the barrier to entry was still too high. HOT had been building a teaching curriculum called LearnOSM.org, but in our opinion it was still not approachable for the non-geo person and that there needed to be better path to on-boarding new volunteers. So the first goal of MapGive is to make it easier for new mappers to get involved by teaching them why humanitarian mapping and open data are important, how to map in OSM, how to use the Tasking Manager, and then connecting them directly with vetted humanitarian mapping projects. In a similar fashion as we had done with satellite imagery, we needed to leverage resources that the Department had that could fill gaps that the HOT community did not do well.

MapGive's second goal was to build an outreach strategy utilizing social media to recruit new mapping volunteers, and State Department has an amazing pool of candidates to recruit from. With approximately 30 million fans and followers on social media around the world, the Department has curated a community that want to interact with America. After investing in building this community, the obvious next question was to determine if could be motivated to action. As I stated at the time, our challenge was to transform social media followers into social media do-ers. Additionally, each regional bureau and embassy has its own public affairs section that conducts outreach with the local population. With MapGive we wanted to give the diplomats working in-country a capability to engage their local communities around the idea of humanitarian mapping and host a mapathon for themselves.

Third, I had quickly realized that IttC needed a communication strategy and support engaging with the media, and that we needed to leverage the incredible skills of our public diplomacy colleagues in the Department who specialized in this work. OIE brought a team of specialists in web design, user interface design, multimedia production, communications, and social media strategy – who were all non-mappers – together on the project. This group was essentially the equivalent of an in-house design team and they provided a unique opportunity to have a collective set of fresh eyes on the challenges of getting new mappers into the process, and how to craft a communications message to go with it. MapGive needed to ease the communication burden off the HIU and at the same time increase its engagement with the media, a challenge that has only grown as media attention on humanitarian mapping increases.

The strategy for accomplishing these three goals involved rebranding IttC to be more representative of these broader themes and to build a specific web site that would become the central resource for the instructional videos and stories

Educational Campaign (Website)

From a technology perspective the MapGive website was designed with modern web principles in mind, responsive CSS design to allow the website to work natively across all devices, optimized for mobile data connections, so the code had to be fast and small, embedded analytics, and a focus on user experience to make the navigation and layout intuitive. The design is built to be Americans with Disabilities Act / Section 508 compliant which means screen readers and other accessibility technology is fully integrated. This design model also makes it easier to translate the site into other languages. All of the code for the website is open source and available on the HIU GitHub account. Within a couple weeks of the MapGive launch, a group in Italy cloned the code from GitHub, translated the site materials and video captions, and republished the site at: http://mapgive.openstreetmap.it/

The hallmark features of the site are the three videos on Why Map, Learn to Map in OSM, and Learn how to use the Tasking Manager. The videos have a very high production value and are fully captioned which again simplifies the process of translating them into other languages. The site also has sections for user stories, information about open mapping, an FAQ, and instructions on how to host a mapping event.

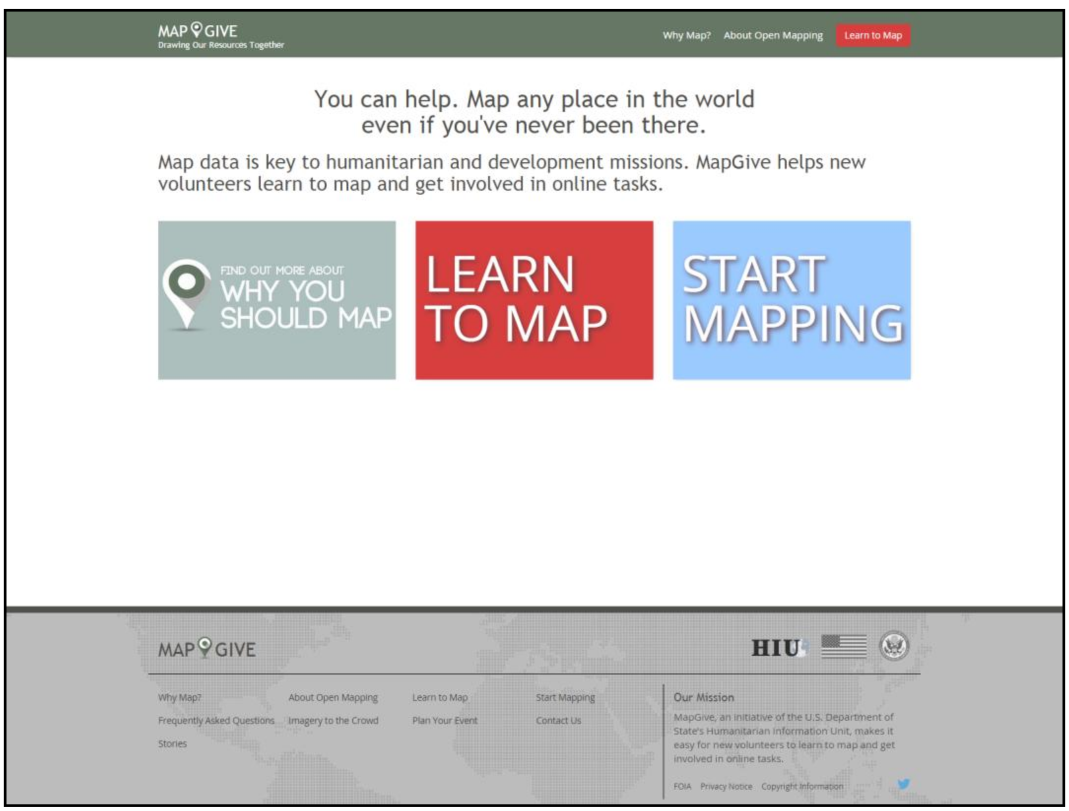

Figure 59 – MapGive homepage

Figure 60 – MapGive homepage, translated into Italian

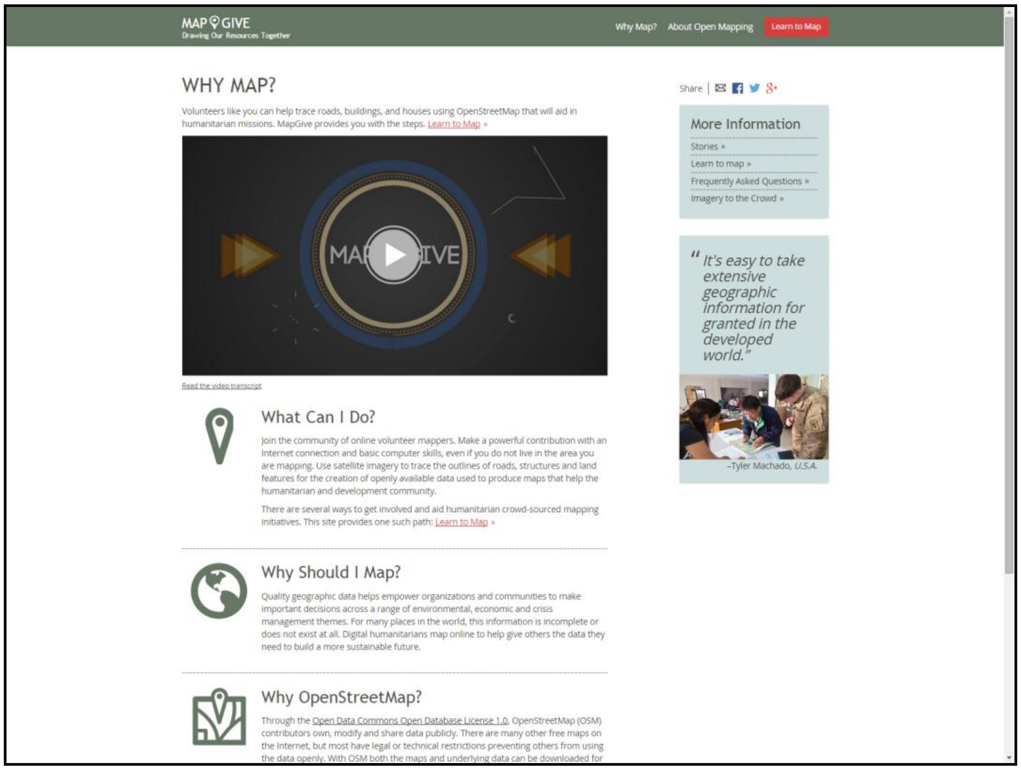

Figure 61 – Why Map? page on MapGive

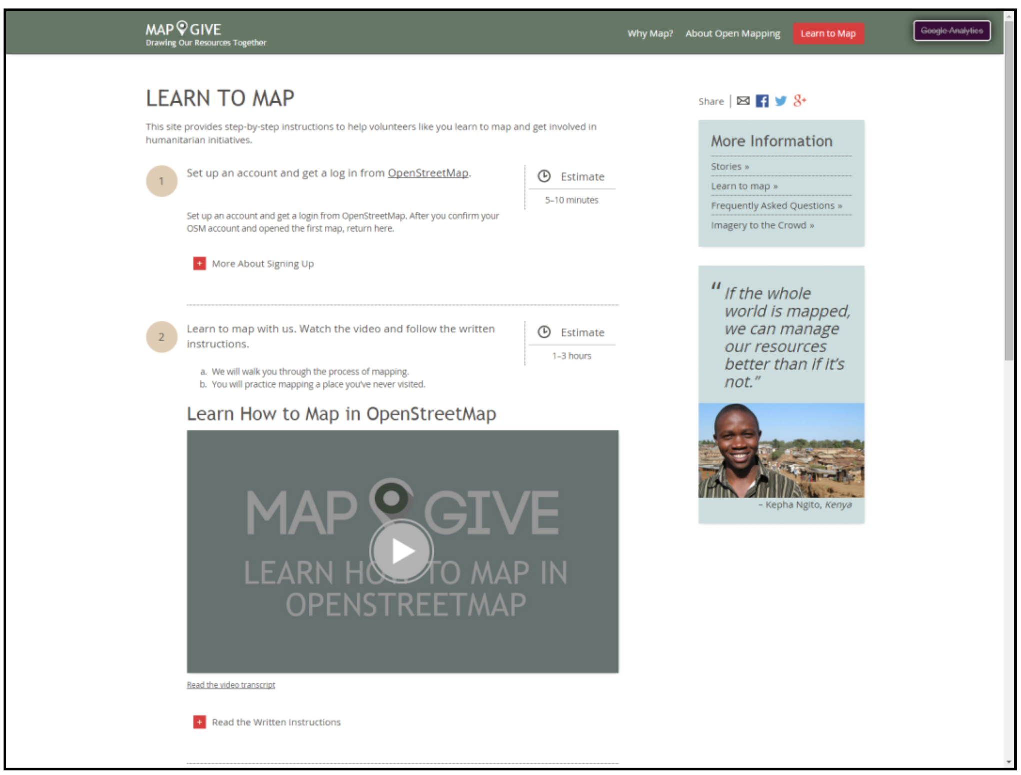

Figure 62 – Learn to Map page on MapGive website with video tutorial about OpenStreetMap mapping

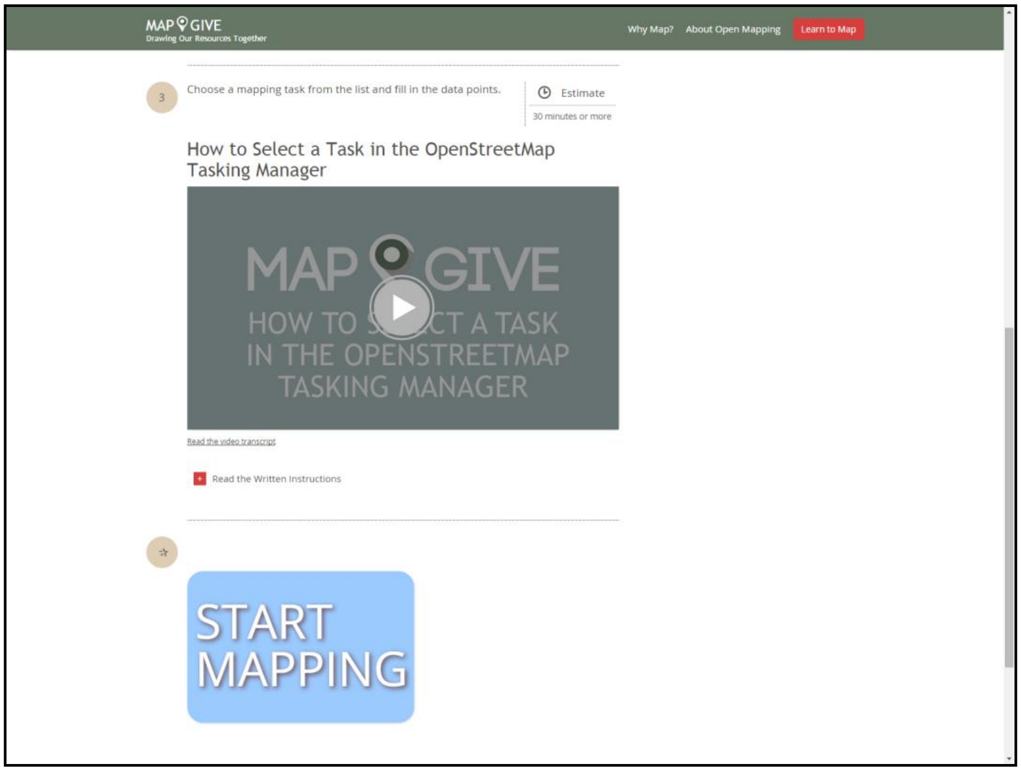

Figure 63 – Learn to Map page on MapGive website with video tutorial about OSM Tasking Manager

Recruitment Campaign (Social Media)

From the social media perspective, a complete Social Media Toolkit (Appendix 5) was created as well as a set of graphics, images, and videos that could be included as “sharables” in social media posts, Figures X and Y. The toolkit was designed specifically with communications professionals (both internal and external) in mind, and was distributed to all the social media coordinators in the Department. We also wanted to leverage the State Department blog, DipNote as a tool for publishing content. So far three MapGive blog posts have been published, one at the launch in March 2014, one at the one year anniversary, and one on the recent Nepal response (Figures 66, 67, 68).

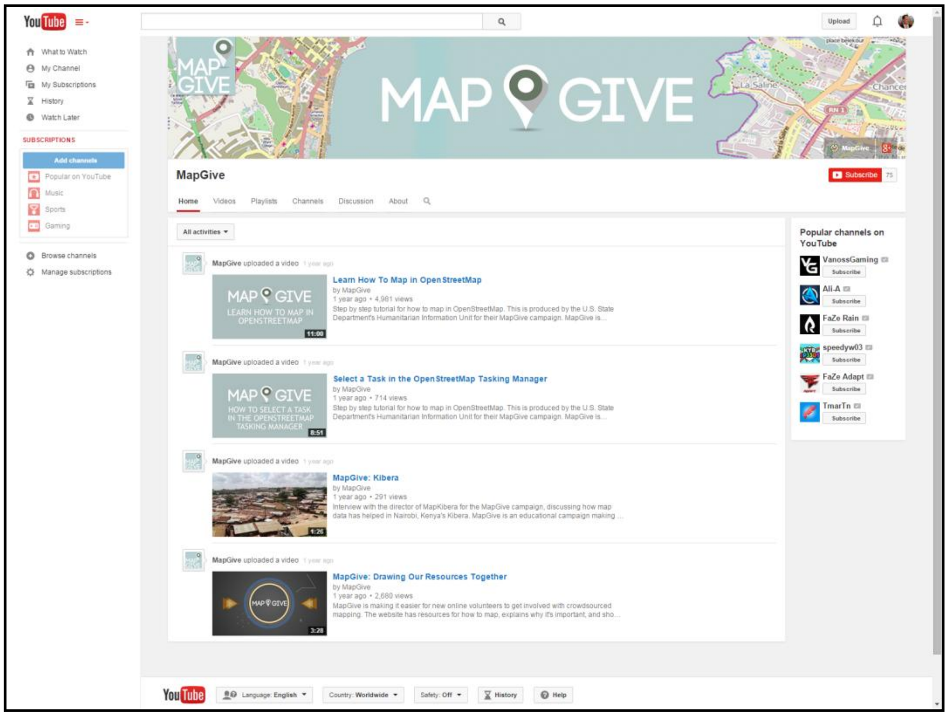

Figure 64 – MapGive YouTube channel

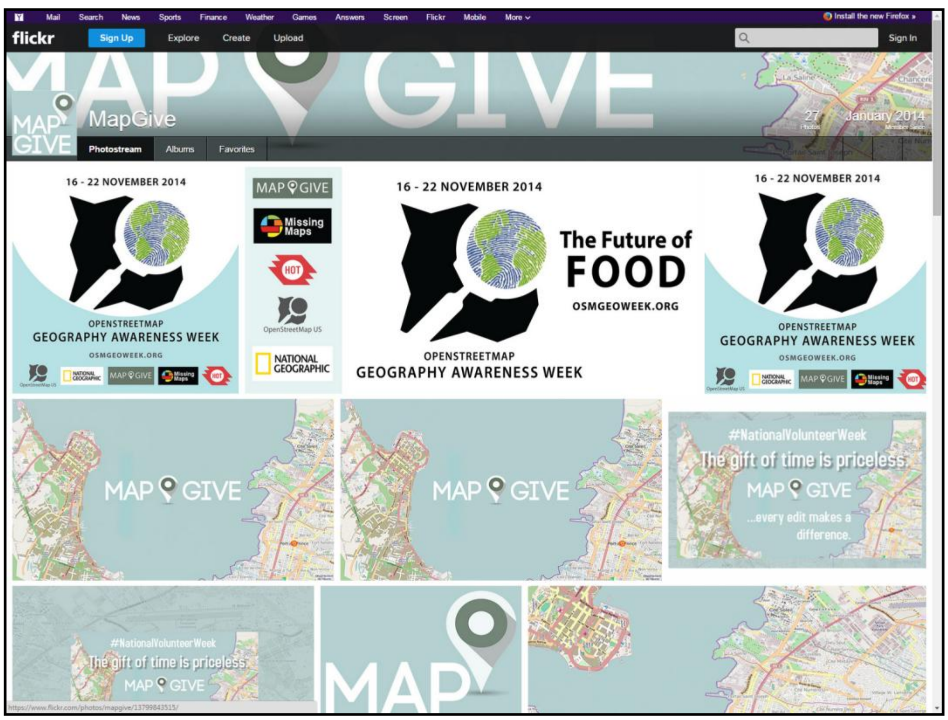

Figure 65 – MapGive Flickr site

Figure 66 – MapGive launch blog, March 7, 2014 (Campbell 2014)

Figure 67 – MapGive One Year Summary, February 21, 2015 (Wilder and Maron 2015a)

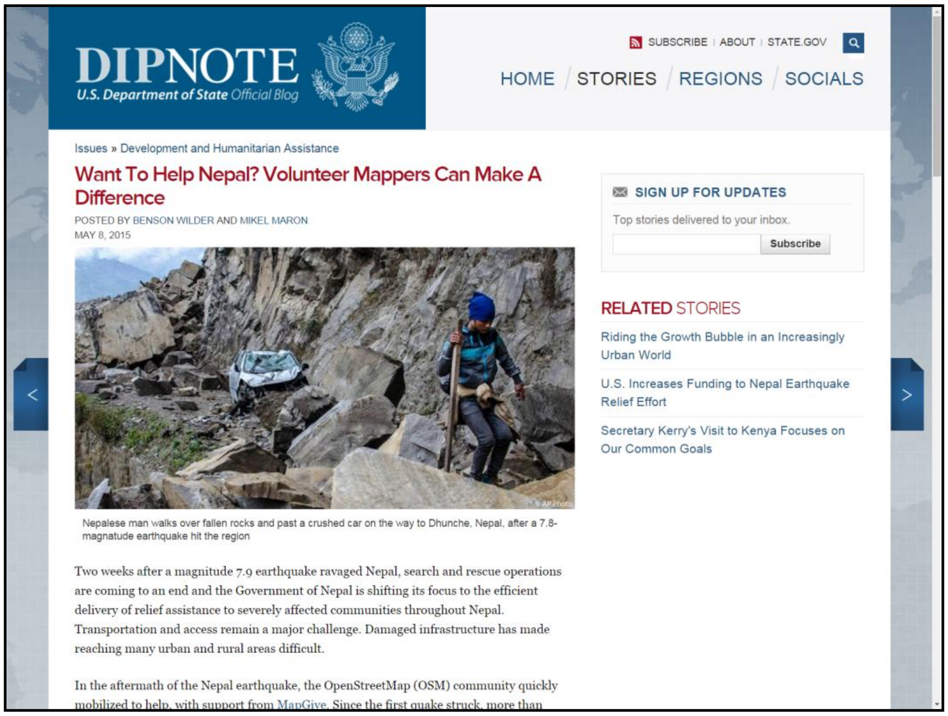

Figure 68 – MapGive Outreach for Nepal Earthquake (Wilder and Maron 2015b)

Mapping Parties

As part of the outreach strategy, MapGive has begun sponsoring mapping parties with universities and NGOs. These have proven to be an effective way to coordinate action around a given task and recruit potential new volunteers. And while there a lot of good stories that emerge from these gatherings, I want to conclude with one in particular. The American Consulate General in Jerusalem sponsored a mapathon that connected 40 Palestinian college students in Ramallah with American counterparts in New York for a six-hour mapping session. The mapping tasks were for the Ebola affected region of West Africa. The picture in Figure 69 is what this project is all about: American diplomacy conducting outreach to teach Palestinian students how to map in OSM, so they can help support a pandemic response effort continent away. This is a scenario that the State Department can replicate a thousand different ways, and potentially make a huge difference.

Figure 69 – Ebola mapathon in Ramallah (Consulate General of the United States Jerusalem 2014)

Future Directions

The accomplishments of the Humanitarian OpenStreetMap Team are impressive, but there are several areas of improvement in the volunteer mapper workflow that must be addressed in order to maintain volunteer enthusiasm for the project and ensure humanitarian organizations continue to use the data. Some of the issues here can be solved with technology, others with organizational changes.

Data Verification and User Ranking

Benkler and Nissenbaum (2006) indicate that peer production systems must have methods to verify / validate data and assign credibility to users. In the current instance of the OSM Tasking Manager, validation is an available feature, but for a variety of reasons, its implementation suffers. The process is manual in that a second mapper, acting in the role of validator, visually reviews a completed tile and determines if the work is acceptable. If yes, the tile is marked as validated and the visual display of that tile changes color. If not, the validator should inform via the comment box what is wrong and mark the tile incomplete, returning it to the pool of available tiles. Technologically the Tasking Manager can accept a “whitelisted” set of users that can act as validators, but the feature is rarely, if ever, used. The problem is organizational, there is no established process for determining who can be a validator and build the list of users.

What ends up happening is that usually one or two advanced mappers take it upon themselves to act as validators and review all, or a subset, the tiles in a task. This process can work in a small activation of limited scope and number of mappers, but it quickly breaks down in large mapping events or disasters when the number of new mappers is high. In these cases you often get new volunteers marking tiles as validated, when they are barely mapped. Numerous instances of this were observed during MapGive mapathons. The problem is two-fold, new volunteers do not understand the validation system and the existing controls are not used to stop them from making the mistake. While the problem is easy to diagnose, the solution is tougher to implement. If the process is to remain manual, using the same system that in place, then the HOT activation community must determine a list of approved validators, and then ensure those users remain committed to the task. Supplemental educational material should also be created to help educate new volunteers about the process.

I believe the solution to this problem is technological, and requires the development of new capabilities and applications. The converse problem of data validation is user rankings as there is no system to evaluate a user's quality of mapping. A technological solution could address both of these problems. The solution to the issue is to bring more volunteers into the validation process, but do so in a way that does not require them to become mappers. This approach has several advantages, first it harkens back to Eric Raymond's famous quote (2001) regarding the security of open source software, "given enough eyeballs, all bugs are shallow." A notion that was later enshrined an Linus's Law, in honor of Linux inventor Linus Torvalds, and the spatial corollary tested in OSM by Haklay et al. (2010). Second, it lowers the bar for participation in HOT and offers new volunteers a less intimidating path to supporting mapping tasks.

I propose the development of a mobile application that is linked with the Tasking Manager so that when a volunteer completes a map tile a validation process is set in motion. The idea is to sub-divide a completed tile into a number of smaller subset, perhaps 100 individual tiles. The sub-tiles should be of a dimension that fits on a mobile screen and displays a limited amount of mapped vector data displayed on top of the imagery used to map from. The application user would then visually review the sub-tile and mark it as acceptable or not. Over time that same sub-tile would be displayed to several different users and their rating recorded. For instance, if four out of five reviewers approved the sub-tile, it is likely correct. In this way it is possible to crowdsource the validation and to integrate that data back into the Tasking Manager via a "validation surface" that would correspond geographically to the area being mapped. Alternatively, if it is receiving a negative rating then it should automatically trigger the Tasking Manager to mark the tile as incomplete and flag an advanced mapper to review the work.

An approach like this would capture metrics about several users simultaneously, who mapped the data originally and who reviewed the sub-tile. As more reviews were completed it would be possible to evaluate the quality of each users work and begin assigning an empirically derived confidence score to them. The system would provide feedback to users as there would be a record of what they mapped, how it was reviewed, if it was modified, and what changes were made. This way a user could learn from the validation process, a feedback loop that can keep volunteers engaged and improving. A submission to the HOT email list serve during the recent Nepal earthquake response started a long discussion of these issues, reporting frustration with both the lack of education of new volunteers and the lack of feedback in the mapping process (Aytoun 2015). Additionally, the need for the development of trust factors or confidence scores was highlighted as the concluding thought of Neis and Zielstra (2014, 96) thorough review of OSM.

User Identification and Pairing

Benkler (2002) highlights the matching of volunteer talent with the appropriate task as one of the advantages of commons-based peer production systems over less-efficient models. Unfortunately, however, this capability does not exist in OSM, HOT, or the OSM Tasking Manager. As stated above, the lack of data validation systems currently precludes any automated assessment of user quality. And without the ability to know a user's skill level, it is impossible to direct them to an appropriate task in any automated way. On occasion a mapping task in the Tasking Manager will indicate it is for advanced mappers only in the Instructions text box, but the system will still allow any user to access the task.

In order to facilitate the appropriate pairing of user and task, two things are needed. First is the confidence score already discussed, and second is a rating of the complexity of a map tile. From personal mapping experience, the complexity of any given map tile can vary tremendously. Rural South Sudan is significantly easier to map than downtown Kathmandu or a typhoon wrecked Tacloban, Philippines. What is needed is a methodology for assessing the mapping complexity for each tile. Significant improvements in this process could be gained by a quick, human interpretation of the complexity of a map tile and assigning it a qualitative score. The Tasking Manager could represent this score via a color ramp and provide guidance about the skill level and amount of time needed to complete the task in the Instructions section describing the task.

An automated approach is more complex, with an optimal system using some form of remote sensing classification of each tile combined with an analysis of the existing OSM data in the tile and qualitative feedback from users after they work on it. In terms of the remote sensing function, I believe some type of texture measure using variance or entropy would be a good starting point. Regardless of the exact function, over time it would be calibrated with feedback from the user surveys.

Automated Feature Extraction: Combined Human – Computer Approaches

The reason I choose crowdsourcing as the preferred method of image interpretation is skepticism in the utility of automated approaches to vector extraction from imagery, particularly as the objects being mapped occur at finer spatial scales. And while I still believe this, it is hard not to look at OSM as the world's largest remote sensing training database. One research direction that deserves some attention is a workflow that would pre-process image tiles, likely using image segmentation algorithms, and compare the outputs to data mapped into OSM. With recent advancements in machine vision and increasing resolution of satellite imagery, perhaps the mapping process could be shortened if objects were already extracted and tagged, and the human mapper focused less on redundant mapping tasks and more on correcting erroneous extractions.

Analytics

Understanding the amount of change that occurs within a given task remains too difficult. The best option for analytics is the ResultsMap website run by academic and OSM researcher Pascal Neis (Neis 2015). However even this site only catalogs specific changeset level data for seven days, and the software code is not made openly available. The other alternative is to setup a system that downloads the OSM planet file (the entire OSM dataset) or some country level subset at a regular interval and analyze the differences. This is a complex undertaking requiring significant skill to setup. At least in the context of HOT activations, a suite of automated analytics should be developed that make it easy to snapshot the OSM data before, during, and after a task is published, and analyze the change between versions. These can make powerful visualizations that users often find compelling. From a diplomacy perspective, the ability to analyze where in the world edits to a task are coming from would be very useful. Imagine a two paneled map where edits to the OSM basemap are being show in real-time on one panel, and the country they originated in is flashing on the other. Doing this would require some thinking about how to track a user through both the imagery service and the OSM data checkins, but it could be done, and by keeping it at the country level, would not violate privacy concerns.

Data Expiration

Lastly, I have grown more concerned about map data created from IttC or MapGive projects not changing to keep pace with develop over time. Recently I was looking back at the Horn of Africa refugee camps, and in many areas the camps have changes so significantly that the OSM data is now wrong. We need to develop a mechanism to re-evaluate data created as part of tasks that we sponsor. This system could leverage the mobile application presented in the data verification section.

Conclusions

On a personal level, the success of Imagery to the Crowd was a tremendous achievement, and one that took long, hard hours from me and a lot of other people. It brought together 19 months of legal and policy negotiations to be able to share the imagery, with 14 months of technical development on the CyberGIS. More importantly, it proved that my underlying assumptions about peer production networks, geospatial cyberinfrastructure, and the HIU's ability to replicate the Haiti response were correct. It is possible to help catalyze the crowd to produce amazing data. And we did it all using satellite imagery tax-payers had already paid for, powered by free and open source geographic software, deployed in a cloud computing environment that cost a couple dollars an hour to run. And as IttC has grown, the results have only become more astounding.

With the full deployment of the CyberGIS and MapGive, the State Department now has two functional geographic capabilities that have been vetted and approved for use. Collectively these projects provide the software tools, data sources, and methodologies needed to interface with both the foreign policy and public diplomacy elements of the Department. With its global reach, and mission of promoting peace, security, and economic prosperity, the State Department is uniquely positioned to engage, educate, and equip local populations to harness the power of open geographic data and spatial analysis for themselves.

Reference Cited

Anderson, T. J. 2015. Nepal Earthquake. *osm.townsendjennings.com*. http://osm.townsendjennings.com/nepal/ (last accessed 3 May 2015).

Aytoun, R. 2015. AAAAAAAAAAAAAARRRRRRRRGH! *Humanitarian OpenStreetMap Team*. https://lists.openstreetmap.org/pipermail/hot/2015-April/008191.html (last accessed 27 April 2015).

Baldwin, C., and E. von Hippel. 2010. Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation. *Harvard Business School Finance Working Paper No. 10-038; MIT Sloan Research Paper No. 4764-09* :37.

Benkler, Y. 2002. Coase's Penguin, or, Linux and "The Nature of the Firm." *Yale Law Journal* 112 (3):369–446.

Benkler, Y., and H. Nissenbaum. 2006. Commons-based Peer Production and Virtue. *Journal of Political Philosophy* 14 (4):394–419.

Berger, S. 1994. US Policy on Foreign Access to Remote Sensing Space Capabilites. http://nsarchive.gwu.edu/NSAEBB/NSAEBB404/docs/07.pdf (last accessed 8 May 2015).

Brahimi, L. 2000. *Report of the Panel on United Nations Peace Operations*. New York: United Nations.

Brandt, H., and J. S. Campbell. 2013. Crowdsourcing Diplomacy. *SXSW PanelPicker*. http://panelpicker.sxsw.com/vote/21573 (last accessed 29 April 2015).

Bullock, K. 2014. Mapping the World with OpenStreetMap via Mapbox. *DigtialGlobe Blog*. http://www.digitalglobeblog.com/2014/04/09/mappingtheworldinosm/ (last accessed 8 May 2015).

Campbell, J. S. 2014. OpenStreetMap for Diplomacy: MapGive and Presidential Innovation Fellow. *DipNote | U.S. Department of State Official Blog*. http://blogs.state.gov/stories/2014/03/07/openstreetmap-diplomacy-mapgive-and-presidential-innovation-fellow (last accessed 21 April 2015).

Capelo, L., N. Chang, and A. Verity. 2012. *Guidance for Collaborating with Volunteer & Technical Communities*. Digital Humanitarian Network. http://digitalhumanitarians.com/resource/guidance-collaborating-volunteer-technical-communities-english (last accessed 5 May 2013).

CartoDB. 2015. High Resolution Satellite Imagery on Demand. *cartodb.com*. http://cartodb.com/solutions/satellite-aerial-image (last accessed 8 May 2015).

Christensen, C. M., and M. E. Raynor. 2003. *The Innovator's Solution: Creating and Sustaining Successful Growth*. Boston: Harvard Business School Publishing.

Christian, C. 2007. Location, Location, Location... *Foreign Service Journal* 84 (January):37–41.

Clarke, K. C. 1995. *Analytical and Computer Cartography* 2nd ed. Upper Saddle River, NJ: Prentice Hall.

Consulate General of the United States Jerusalem. 2014. American Consulate General Co-Sponsors Palestinian University Students' Efforts to Map West African Nations Stricken with Ebola. *2014 Press Releases*. http://jerusalem.usconsulate.gov/pr-10182014.html (last accessed 28 April 2015).

Cowan, N. 2011. A Geospatial Data Management Framework for Humanitarian Response. In *Proceedings of the 8th International ISCRAM Conference*. Lisbon, Portugal http://www.iscramlive.org/ISCRAM2011/proceedings/papers/295.pdf (last accessed 7 March 2013).

Crampton, J. W., S. M. Roberts, and A. Poorthuis. 2014. The New Political Economy of Geographical Intelligence. *Annals of the Association of American Geographers* 104 (1):196–214.

Crisis Mappers. 2015. CrisisMappers - The Humanitarian Technology Network. *crisismappers.net*. http://crisismappers.net/ (last accessed 8 May 2015).

Crowley, J. 2015. History of NextView Imagery sharing, prototype examples from Afghanistan, role of RELIEF.

Crowley, J., and J. Chan. 2010. Disaster Relief 2.0: The Future of Information Sharing in Humanitarian Emergencies. *Harvard Humanitarian Initiative and UN Foundation-Vodafone Foundation-UNOCHA*.

Currion, P. 2006. Better the Devil We Know: Obstacles and Opportunities in Humanitarian GIS. *humanitarian.info*.

DigitalGlobe, Inc. 2015. *DigitalGlobe 2014 Annual Report*. Washington, D.C: United States Securities and Exchange Commission. http://investor.digitalglobe.com/phoenix.zhtml?c=70788&p=irol-IRHome.

Dobson, J. 2012. Through the Macroscope: Geography's View of the World. *ArcNews* Winter (2011/2012). http://www.esri.com/news/arcnews/winter1112articles/through-the-macroscope-geographys-view-of-the-world.html (last accessed 7 May 2015).

Firebaugh, J. 2013. New Map Editor Launches on OpenStreetMap.org. *Mapbox.com*. https://www.mapbox.com/blog/new-map-editor-launches-openstreetmap/ (last accessed 14 April 2015).

Free Software Foundation. 2007. GNU General Public License v3.0. *GNU Project*. https://gnu.org/licenses/gpl.html (last accessed 6 May 2015).

———. 2015. What is free software? *GNU Project*. https://www.gnu.org/philosophy/free-sw.html (last accessed 13 April 2015).

GitHub. 2015. Features. https://github.com/features (last accessed 5 May 2015).

Goodchild, M. F. 2007. Citizens as Sensors: The World of Volunteered Geography. *GeoJournal* 69:211–221.

Haklay, M. 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design* 37 (4):682–703.

Haklay, M. (Muki), S. Basiouka, V. Antoniou, and A. Ather. 2010. How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. *The Cartographic Journal* 47 (4):315–322.

Haklay, M., A. Singleton, and C. Parker. 2008. Web mapping 2.0: The neogeography of the GeoWeb. *Geography Compass* 2 (6):2011–2039.

Von Hippel, E. 1986. Lead Users: A Source of Novel Product Concepts. *Management Science* 32 (7):791–805.

Humanitarian Information Unit. 2012. Horn of Africa Refugee Database. https://hiu.state.gov/Data/HornofAfrica_RefugeeDatabase_2012January_HIU_USDoS.zip (last accessed 25 April 2015).

———. 2015a. Humanitarian Information Unit - Data. *HIU Data*. https://hiu.state.gov/data/data.aspx (last accessed 8 May 2015).

———. 2015b. Humanitarian Information Unit GitHub Account. *github.com*. https://github.com/state-hiu (last accessed 3 May 2015).

Humanitarian OpenStreetMap Team. 2015a. About the OSM Tasking Manager. *OSM Tasking Manager*. http://tasks.hotosm.org/about (last accessed 3 May 2015).

———. 2015b. Humanitarian OpenStreetMap Team. http://hotosm.org/ (last accessed 2 May 2015).

Kralidis, A. T. 2008. Geospatial Open Source and Open Standards Convergences. In *Open Source Approaches in Spatial Data Handling*, eds. G. B. Hall and ·. M. G. . Leahy, 1–20. Berlin: Springer.

Kuhn, T. 1996. *The Structure of Scientific Revolutions* 3rd ed. Chicago: University of Chicago Press.

Liu, S. B. 2014. Crisis Crowdsourcing Framework: Designing Strategic Configurations of Crowdsourcing for the Emergency Management Domain. *Computer Supported Cooperative Work (CSCW)* 23 (4-6):389–443.

Liu, S. B., and J. Ziemke. 2013. From cultures of participation to the rise of crisis mapping in a networked world. In *The Participatory Cultures Handbook*, eds. A. Delwiche and J. J. Henderson, 185–196. New York: Taylor & Francis.

MapBox. How do web maps work? *Mapbox.com*. https://www.mapbox.com/guides/how-web-maps-work/ (last accessed 25 April 2015).

Maron, M. 2007. Yahoo approves of JOSM plugin. *OSM-Talk*. https://lists.openstreetmap.org/pipermail/talk/2007-July/015682.html (last accessed 4 April 2015).

Meier, P. 2012. Crisis Mapping in Action: How Open Source Software and Global Volunteer Networks Are Changing the World, One Map at a Time. *Journal of Map & Geography Libraries* 8 (2):89–100.

Meier, P., and R. Munro. 2010. The unprecedented role of SMS in disaster response: Learning from Haiti. *SAIS Review of International Affairs* 30 (2):91–103.

MetaCarta Labs. 2010. TileCache. *TileCache.org*. http://tilecache.org/ (last accessed 25 April 2015).

National Research Council. 2002. *Down to earth: geographic information for sustainable development in Africa*. Washington, D.C: National Academy Press.

———. 2007. *Successful response starts with a map: improving geospatial support for disaster management*. Washington, D.C: National Academies Press.

National Science Foundation Cyberinfrastructure Council. 2007. *Cyberinfrastructure Vision for the 21st Century*. Arlington, VA: National Science Foundation. http://www.nsf.gov/pubs/2007/nsf0728/index.jsp.

Naval Postgraduate School. 2011. RELIEF 11-4. *RELIEF Newsletter* (2). http://star-tides.net/sites/default/files/documents/files/RELIEF%2011-04%20Newsletter.pdf (last accessed 13 April 2015).

Neis, P. 2015. Overview of ResultMaps. *neis-one.org*. http://resultmaps.neis-one.org/ (last accessed 2 May 2015).

Neis, P., and D. Zielstra. 2014. Recent Developments and Future Trends in Volunteered Geographic Information Research: The Case of OpenStreetMap. *Future Internet* 6 (1):76–106.

Okolloh, O. 2009. Ushahidi, or "testimony": Web 2.0 tools for crowdsourcing crisis information. *Participatory Learning in Action* 59 (Change at Hand: Web 2.0 for Development):65–70.

Open Data Commons. ODC Open Database License (ODbL) Summary. *opendatacommons.org*. http://opendatacommons.org/licenses/odbl/summary/ (last accessed 7 May 2015).

Open Geospatial Consortium. 2007. Simple Feature Access - Part 2: SQL Option. http://www.opengeospatial.org/standards/sfs (last accessed 8 May 2015).

OpenLayers. 2015. OpenLayers 3. *OpenLayers.org*. http://openlayers.org/ (last accessed 25 April 2015).

Open Source Geospatial Foundation. 2012. Tile Map Service Specification. *OSGeo Wiki*. http://wiki.osgeo.org/wiki/Tile_Map_Service_Specification (last accessed 25 April 2015).

OpenStreetMap Foundation. 2014. OpenStreetMap Foundation Wiki. http://wiki.osmfoundation.org/wiki/Main_Page (last accessed 7 May 2015).

OpenStreetMap legal-talk Archives. https://lists.openstreetmap.org/pipermail/legal-talk/ (last accessed 9 April 2015).

OpenStreetMap Wiki. 2015a. Bing Imagery. *OpenStreetMap Wiki*. http://wiki.openstreetmap.org/wiki/Bing (last accessed 5 April 2015).

———. 2015b. Humanitarian OSM Team. *OpenStreetMap Wiki*. http://wiki.openstreetmap.org/wiki/Humanitarian_OSM_Team (last accessed 5 April 2015).

———. 2015c. OpenStreetMap Statistics. *OpenStreetMap Wiki*. http://wiki.openstreetmap.org/wiki/Stats (last accessed 14 April 2015).

———. 2015d. Slippy map tilenames. *OpenStreetMap Wiki*. http://wiki.openstreetmap.org/wiki/Slippy_map_tilenames#Resolution_and_Scale (last accessed 25 April 2015).

———. 2015e. What is OpenStreetMap? *OpenStreetMap Wiki*. http://wiki.openstreetmap.org/wiki/What_is_OpenStreetMap%3F (last accessed 3 May 2015).

O'Reilly, T. 2005. The Open Source Paradigm Shift. In *Perspectives on Free and Open Source Software*, eds. J. Feller, B. Fitzgerald, S. A. Hissam, and K. R. Lakhani, 461–481. Cambridge: MIT Press.

———. 2007. What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *Communications & Strategies* 1:17.

Oros, C. L. 2014. *A brief history of the NPS Field Experimentation program: Spanning STAN, TNT and JIFX*. Monterray, CA: Naval Postgraduate School. http://calhoun.nps.edu/handle/10945/43732.

Perens, B. 1999. The Open Source Definition. In *Open sources: voices from the open source revolution*, eds. C. DiBona, S. Ockman, and M. Stone. Beijing ; Sebastopol, CA: O'Reilly.

PostGIS Project. PostGIS 2.0 Manual. http://postgis.net/docs/manual-2.0/ (last accessed 8 May 2015).

Ramsey, P. 2014. Spatial IT and the Enterprise. *cleverelephant.ca*. http://blog.cleverelephant.ca/2014/11/spatial-it-and-enterprise.html (last accessed 6 May 2015).

Raymond, E. S. 1999. *The cathedral and the bazaar: musings on Linux and Open Source by an accidental revolutionary* Rev. ed. Beijing ; Cambridge, Mass: O'Reilly.

ReliefWeb. 2012. Horn of Africa Crisis: 2011-2012. *ReliefWeb*. http://reliefweb.int/disaster/dr-2011-000029-ken (last accessed 25 April 2015).

Roberts, S. 2014. Open Cities and Crisis Mapping. *USAID Impact*. http://blog.usaid.gov/2014/01/open-cities-and-crisis-mapping/ (last accessed 21 April 2015).

Saveri, A., H. Rheingold, and K. Vian. 2005. *Technologies of Cooperation*. Palo Alto: Institute for the Future. http://www.iftf.org/our-work/global-landscape/cooperation/technologies-of-cooperation/.

Shirky, C. 2011. *Cognitive surplus : how technology makes consumers into collaborators*. New York: Penguin Books.

Sui, D. Z. 2008. The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Computers, Environment and Urban Systems* 32 (1):1–5.

The Open Source Definition. 2015. *Open Source Initiative*. http://opensource.org/osd (last accessed 13 April 2015).

Turner, A. 2006. *Introduction to Neogeography*. Sebastopol: O'Reilly Media.

U.S. Department of State. 2013. 1 FAM 430 Bureau of Intelligence and Research. http://www.state.gov/documents/organization/84193.pdf.

———. 2015a. Humanitarian Information Unit. *Humanitarian Information Unit*. https://hiu.state.gov/Pages/Home.aspx (last accessed 8 May 2015).

———. 2015b. MapGive | Homepage. *mapgive.state.gov*. http://mapgive.state.gov/ (last accessed 2 May 2015).

———. 2014. *Open Government Plan*. Washington, D.C: U.S. Department of State. http://www.state.gov/open/230791.htm.

———. 2015c. *Quadrennial Diplomacy and Development Review: Enduring Leadership in a Dynamic World*. Washington, D.C: U.S. Department of State. http://www.state.gov/s/dmr/qddr/.

Verjee, F. 2007. An assessment of the utility of GIS-based analysis to support the coordination of humanitarian assistance. http://pqdtopen.proquest.com/#viewpdf?dispub=3297449 (last accessed 7 March 2013).

———. 2005. The Application of Geomatics in Complex Humanitarian Emergencies | The Journal of Humanitarian Assistance. *The Journal of Humanitarian Assistance*. http://sites.tufts.edu/jha/files/2011/04/a179.pdf (last accessed 7 March 2013).

Wang, S. 2010. A CyberGIS Framework for the Synthesis of Cyberinfrastructure, GIS, and Spatial Analysis. *Annals of the Association of American Geographers* 100 (3):535–557.

Wang, S., L. Anselin, B. Bhaduri, C. Crosby, M. F. Goodchild, Y. Liu, and T. L. Nyerges. 2013. CyberGIS software: a synthetic review and integration roadmap. *International Journal of Geographical Information Science* 27 (11):2122–2145.

Weiss, T. G., and P. J. Hoffman. 2007. The Fog of Humanitarianism: Collective Action Problems and Learning-Challenged Organizations. *Journal of Intervention and Statebuilding* 1 (1):47–65.

Wikipedia. 2015a. Tiled web map. *Wikipedia*. http://en.wikipedia.org/wiki/Tiled_web_map (last accessed 25 April 2015).

———. 2015b. Wikipedia: Statistics. http://en.wikipedia.org/wiki/Wikipedia:Statistics (last accessed 5 May 2015).

Wilder, B., and M. Maron. 2015a. Open Data Day: How the State Department is Linking Diplomacy With Collaborative Mapping During Crises. *DipNote | U.S. Department of State Official Blog*. http://blogs.state.gov/stories/2015/02/21/open-data-day-how-state-department-linking-diplomacy-collaborative-mapping-during (last accessed 21 April 2015).

———. 2015b. Want To Help Nepal? Volunteer Mappers Can Make A Difference. *DipNote | U.S. Department of State Official Blog*. http://blogs.state.gov/stories/2015/05/08/want-help-nepal-volunteer-mappers-can-make-difference (last accessed 9 May 2015).

Wood, W. B. 2000. Complex emergency response planning and coordination: Potential GIS applications. *Geopolitics* 5 (1):19–36.

World Bank. 2014. *Understanding Disaster Risk in an Evolving World*. Washington, D.C: World Bank.

World Bank, and Humanitarian OpenStreetMap Team. 2014. *Open Data For Resilience Initiative: Planning An Open Cities Mapping Project*. Washington, D.C: World Bank. http://gfdrr.org/sites/gfdrr/files/publication/Planning-an-Open-Cities-Mapping-Project_0.pdf.

Yahoo! Aerial Imagery. *OpenStreetMap Wiki*. http://wiki.openstreetmap.org/wiki/Yahoo!_Aerial_Imagery (last accessed 5 April 2015).

Yang, C., R. Raskin, M. Goodchild, and M. Gahegan. 2010. Geospatial Cyberinfrastructure: Past, present and future. *Computers, Environment and Urban Systems* 34:264–277.

Ziemke, J. 2012. Crisis Mapping: The Construction of a New Interdisciplinary Field? *Journal of Map & Geography Libraries* 8 (2):101–117.

Zook, M., M. Graham, T. Shelton, and S. Gorman. 2010. Volunteered Geographic Information and Crowdsourcing Disaster Relief: A Case Study of the Haitian Earthquake. *World Medical & Health Policy* 2:27.

Appendix 1 – NextView End User License Agreement

NEXTVIEW IMAGERY END USER LICENSE AGREEMENT

1. Introduction. This End User License Agreement ("EULA") is between DigitalGlobe, Inc., a Delaware Corporation ("DigitalGlobe" or "Seller") and National Geospatial-Intelligence Agency ("NGA"), the purchaser of this EULA, which governs the use of the data products or documentation ("Products") accompanying this EULA in accordance with Contract NMA 301-03-3-0001 (the "Contract").

2. Applicability. This license applies to imagery and products licensed under the Contract, including data downlinked to domestic and foreign ground stations.

3. License Granted and Permitted Uses.

a. General Terms

1. This clause applies to all unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data licensed under this Contract. No other clauses related to intellectual property or data rights of any sort shall have any effect related to the unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data delivered under this Contract.

2. All license rights for use of the unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data provided to the U.S. Government purchased under this NGA contract are in perpetuity.

3. Licensed users may generate an unlimited number of hardcopies and softcopies of the unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data for their use.

4. (i) Licensed users may generate any derived product from the licensed unprocessed sensor data; and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data.

(ii) Unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data licensed under this NGA contract have no restrictions on use and distribution, but shall contain the copyright markings.

b. Licensed Users

1. The imagery may be used by the U.S. Government (including, all branches, departments, agencies, and offices).

2. The U.S. Government may provide the imagery to the following organizations:
State Governments
Local Governments
Foreign Governments and inter-governmental organizations
NGO's and other non-profit organizations

3. In consideration for the flexibility afforded to the U.S. Government by allowing unprocessed sensor data and requirements-compliant processed imagery, imagery services, imagery-derived products and imagery support data to be shared, the United States Government shall use its reasonable best efforts to minimize the effects on commercial sales. Acquisition and dissemination of imagery and imagery products collected within the United States shall be restricted in accordance with law and regulation.

DigitalGlobe, Inc.
NextView EULA_5749 Rev 1.0
08/10/05

Appendix 2 – EnhancedView Contract, Section H - Special Contract Requirements

H.22 **(U) Emergencies, Disasters, And Humanitarian Efforts**

(a) (U) In support of emergencies, disasters, and humanitarian efforts, the NGA may disseminate and/or post on open web sites imagery licensed under this contract regardless of whether the recipients are within the NextView license user groups. The imagery will contain the copyright notice and the NextView license notice. After 30 days, the imagery will be handled in accordance with the NextView license.

(b) (U) The contractor will be given notice within 24 hours after the start of the dissemination/posting of imagery under the authority of this clause.

(c) (U) If the contractor does not believe the situation constitutes an emergency, disaster, or humanitarian effort, the contractor has 24 hours after receiving notice to object to the dissemination/posting of the imagery under the authority of this clause. If the parties cannot reach agreement, the matter will be resolved in accordance with the Disputes Clause and the other terms and conditions of this contract.

Appendix 3 – NextView License Acknowledgement

NEXTVIEW LICENSE ACKNOWLEDGEMENT

These data are licensed for use by the US Government (USG) under the NextView (NV) license and copyrighted by COMPANY NAME/DATE. The NV license allows the USG to share the imagery and Literal Imagery Derived Products (LIDP) with entities outside the USG when that entity is working directly with the USG, for the USG, or in a manner that is directly beneficial to the USG. The party receiving the data can only use the imagery or IDP for the original purpose or only as otherwise agreed to by the USG. The party receiving the data cannot share the imagery or IDP with a third party without express permission from the USG. At no time should this imagery or IDP be used for other than USG-related purposes and must not be used for commercial gain. The copyright information should be maintained at all times. Your acceptance of these license terms is implied by your use.

Yes, requester agrees to abide by this statement (documentation attached). ☐

No, requester does not agree to abide by this statement (documentation attached). ☐

Appendix 4 – NextView and You (Original)

NextView License and You

You must:

Properly attribute (mark) all NV Imagery and imagery derived products (IDP) with its Copyright information and educate anyone that receives the data on the license terms

You must not:

Provide/share imagery or IDPs made from NV licensed imagery with anyone planning to sell it or used it for commercial gain

You may:

- Share imagery or IDPs with anyone directly working with/for the USG, including:
  - U.S. Government Employees/Contractors*
  - Universities supporting USG via contract(s)
  - State/Local Governments
  - Foreign Governments
  - Intergovernmental Agencies
  - NGO's & Non-profit Organizations
- Post properly attributed "dumb" IDPs on public web sites
- Post/disseminate imagery using *access-controlled* web/FTE sites

You should seek clarification (see POC list) before:

- Publically releasing or openly disseminating imagery or IDPs with image metadata
- Sharing with Educational Institutions for strictly educational/research purposes (even with USG grants)
- Sharing with a company or other entity that might profit from the imagery shared
- Posting imagery to a web site without access controls
- Allowing Imagery or IDPs to be shared with a third party
- Sharing Imagery or IDPs with Universities with USG grant(s)
- Contacting the Vendors directly

** Contractors' Government sponsor must provide oversight and approval for this sharing arrangement.*

Definitions

- Imagery is the image and associated metadata. Imagery can be further manipulated, enhanced, & processed. Example: GeoPDF, GeoTiff, NITF.
- Image Derived Product (IDP) – any product created from raw imagery – could include image metadata, but generally does not and often referred to as "dumb-image"
- Non-literal Imagery Derived Product is a product derived from the imagery but no longer looks like an image. Example: Line-drawings, maps.
- Third Party Partner – Party otherwise affiliated with the original USG sharing partner, but not the USG directly

Appendix 5 – “MapGive” Campaign: Social Media Toolkit

Table of Contents

Introduction

The Humanitarian Information Unit (HIU) is engaging potential new volunteer mappers by explaining the value of creating open map data and how easy it is to contribute. The “MapGive” campaign will launch on **Friday, March 7**. MapGive aims to increase the number of mapping volunteers by lowering the barriers for entry to mapping. Through the website mapgive.state.gov, MapGive provides training to learn how to map and directs volunteers to mapping tasks with imagery provided by the HIU’s Imagery to the Crowd project (IttC). Additionally, this campaign will serve as a case study for how to turn an international social media audience into a global community of activists.

We’re asking you to help us spread the word and grow the community of volunteer mappers! Please consider sharing the messages below with your social media network.

For questions or concerns, please contact the Humanitarian Information Unit.

Campaign Website

Please direct all traffic to the MapGive website: MapGive.state.gov

Hashtag

Please use the following hashtag in *every* post: #MapGive

Accounts to follow

@MapGive

Topline Messages

- Mapping is important for humanitarian and development efforts, not just in a disaster.
- Anyone with an internet connection can be a volunteer mapper in OpenStreetMap, and you do NOT need to be located in or have specific knowledge of the place you are mapping.
- With crowdsourcing, every effort makes a big difference. You can contribute 20 minutes or an hour to creating open map data.

General Messages

Twitter

You can help map places you've never been before! Learn to #MapGive http://goo.gl/M93H8l

Digital humanitarians #MapGive online to give others data they need to build a more sustainable future. http://goo.gl/SkKhPk

Join the movement: #MapGive in 3 steps! Anyone with a computer & an Internet connection can help http://goo.gl/twjn4T

Curious about how crowdsourced open mapping works? #MapGive explains http://goo.gl/SkKhPk

What's the difference between satellite imagery, map data, and the map itself? #MapGive explains http://goo.gl/eoxBBj

.@hotosm mappers helped after the 2013 typhoon in the Philippines. Join them. Learn to #MapGive http://goo.gl/twjn4T

Volunteer mappers helped Haiti 4 years ago. Now they #MapGive to help before disasters. Join them: http://goo.gl/SkKhPk

How did Kenya's @mapkibera use @openstreetmap to #MapGive & bring change to the community? http://goo.gl/inELrv

Facebook

#MapGive makes it easy for new volunteers to learn to map. Help give others the data they need to build a more sustainable future. http://goo.gl/SkKhPk

Do you have a computer and an Internet connection? Help create open map data! No special skills are needed. Learn how you can help humanitarian and development efforts through #MapGive. http://goo.gl/twjn4T

What's the difference between satellite imagery, map data, and the map itself? #MapGive explains http://goo.gl/eoxBBj

What's an important cause you've contributed to? The World Bank is leading a crowdsourced mapping effort to help the earthquake-prone Kathmandu Valley in Nepal. With more than 500,000 buildings left to map, they still need your help! Become a digital humanitarian with the #MapGive Campaign. http://goo.gl/yWz2eU

Typhoon Haiyan's destruction in the Philippines continues to affect millions, but volunteers like you are still helping save lives! Read how one volunteer made a difference from thousands of miles away using only his laptop. How could your community benefit from open source mapping? http://goo.gl/wMTggV

Visit #MapGive to learn how a map helps communities, like Kibera in Nairobi, Kenya, develop better safety for women and access to education. http://goo.gl/inELrv

Shareables: Photos and Video

Flickr: MapGive

http://www.flickr.com/photos/mapgive

YouTube:

http://www.youtube.com/mapgive

Additional Resources:

Humanitarian Information Unit's Imagery to the Crowd

Humanitarian OpenStreetMap Team

OpenStreetMap

Appendix 6 – MapGive Launch Email

The Department of State's Humanitarian Information Unit is pleased to announce the launch of MapGive, a campaign that seeks to engage worldwide audiences in crowdsourced mapping.

We're asking you to help us make this a success by visiting mapgive.state.gov and sharing on Twitter and Facebook today. Please feel free to use these sample tweets and Facebook posts, or find others attached in the social media toolkit which your organization can use to promote this campaign:

Find us on Twitter @MapGive and use #MapGive in your tweets and on Facebook!

Twitter -- .@StateDept launches @MapGive helping digital humanitarians create open map data in @openstreetmap http://goo.gl/M93H8l

Facebook -- The US State Department launches #MapGive to help digital humanitarians create open map data in OpenStreetMap. Find @Mapgive on Twitter. Visit the site to learn more http://goo.gl/M93H8l

Many of you are already familiar with the Imagery to the Crowd initiative that 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 volunteers can easily map into OpenStreetMap. Until now, several successful projects have been powered by a niche community of mappers already familiar with GIS and OpenStreetMap. We've all asked the question: How much more good could we do if we were able to unlock the mapping power of a crowd of hundreds, even thousands of people?

Through MapGive, the Department has now increased its capacity to bring new online mappers into the community on projects that can have a major impact on humanitarian and development efforts around the world. We've created a website and a social media toolkit to support an educational campaign that aims to teach people about the importance of mapping, give them the skills to map, and help them get connected with mapping tasks.

The Humanitarian Information Unit partnered with the Bureau of International Information Programs' Office of Innovative Engagement (OIE) on this exciting campaign. OIE's office director, Hilary Brandt, and I will be presenting MapGive for the first time this Sunday, March 9th at the SXSW Interactive Conference in Austin, Texas.

Please contact me directly with questions, and to find out how you can get involved with the project.