Open Data Dreams in a Time-Critical Reality: The Evolution of OpenStreetMap in Response to Humanitarian Events

Authors
Affiliation
Address
e-mail address

ABSTRACT
OpenStreetMap (OSM) is a volunteer-driven, globally distributed organization whose members work to create a common digital map of the world. OSM embraces ideals of open data, and to that end innovates both socially and technically to develop practices and processes for coordinated operation. This paper provides a brief history of OSM and then, through quantitative and qualitative examination of the OSM database and other sites of articulation work, examines changes through the lens of two catastrophes that spurred enormous humanitarian relief responses—the 2010 Haiti Earthquake and the 2013 Typhoon Yolanda. The temporally- and geographically-constrained events scope analysis for what is a rapidly maturing, whole-planet operation. The first disaster identified how OSM could support other organizations responding to the event. However, to achieve this, OSM had to refine mechanisms of collaboration around map creation, which were tested again in Typhoon Yolanda. The transformation of work between these two events yields insights into OSM's longer-terms objectives as an organization.

Author Keywords
Crisis informatics, crowdwork, geospatial data, mapping, open data, social computing

ACM Classification Keywords
H.5.3. Groups & Organization Interfaces—collaborative computing, computer-supported cooperative work; K.4.2. Social Issues

General Terms
Human Factors

INTRODUCTION
OpenStreetMap (OSM) sees itself as a global community that creates a common open digital map of the world. OSM is therefore understood to be both a volunteer-driven organization as well as the map it produces for anyone to use: its members work collaboratively to create a high-resolution, global cartographic representation of the planet. Since its creation in 2004 by computer scientists in the UK who were frustrated by the licensing restrictions of the national mapping agency [5], OSM has expanded to a global organization with over 1.5 million registered users and local groups active in over 80 countries. In tandem with this growth in membership, OSM map data have also grown to serve a wide range of uses, including adoption by government [9]; private sector organizations including Microsoft Bing, Foursquare, and Telenav; and, of most relevance to this paper, humanitarian organizations working on disaster relief and international development [26,27]. OSM has sought to manage this expansion of scale and scope through a series of efforts to attract new members, articulate the work of ever-larger numbers of volunteers, and deliver map data in more accessible and reliable ways to a growing user base.

OSM is often called "the Wikipedia of Maps" because it is a volunteer-driven platform where users collaborate to create freely accessible information. People are motivated to contribute for a variety of reasons but many cite the value of open spatial data as an important factor for their participation [4]. As a geowiki [10], OSM appears to operate similarly on its surface to Wikipedia, though a great deal less is known about OSM's normative structures—its reward systems, the details of its array of collaborative practices to generate and edit data, and the manner in which it sees and conducts itself as an organization. As a subject of research, it has not yet gained the attention that Wikipedia has, and seems to be overlooked as a site of social computing research, perhaps because geography in HCI is relatively new (see Hecht's calls for *geo-HCI* [11]).

However, the study-ability of OSM is also very likely at issue here. That the primary site of collaboration is the map database itself rather than human-readable text significantly complicates analysis. Unlike Wikipedia's articles, the individual objects that comprise the OSM map do not have *Talk* pages that would allow researchers to view the objects' change histories, nor the rationales of the editors who modified them. These factors complicate both the study of cooperative behavior in OSM as well as the community's own ability to organize, as we will discuss.

This paper is under review for CHI 2015.

Furthermore, the form that the data take makes analysis complex: OSM datasets are large and unwieldy, and each geographic entity—even down to the level of a single building for example—is itself a series of *nodes* and *ways* with their own attributes and revision histories. Analyzing user activity and changes to geographical entities is therefore a computationally demanding task with few agreed-upon analytics for studying map data creation and in-map interaction meaningfully.

Overview of Everyday Remote Mapping in OSM

The OSM community is segmented into subgroups that self-organize around geographical areas or themes such as mapping wheelchair accessibility or cycling routes. Mappers can work from anywhere in the world using OSM tools and high-resolution satellite imagery to digitize roads, building footprints, and other identifiable features to add to the database. This work is interdependent in that, although contributors work in a distributed manner, there is only one database in which to represent a finite number of objects in the world. The simplest form of coordination occurs when each feature is traced once; then, when different mappers trace additional portions of the same or a related feature—a long road for instance, or surrounding buildings in a complex—their contributions are connected and internally consistent. Ideally, features are tagged in a fashion consistent with best practices as defined by the community through discussion on the wiki page and mailing lists. Simple coordination like this between mappers does take place, but some back-and-forth editing and overwriting are not unusual. However, the precise degree and nature of interaction is unclear because it is quite hard to meaningfully extract this behavior from the database. Observationally, however, it is clear that interactions "in the map"—the map database—as well as discussions "off the map"—email, chats, conferences— are important to OSM's evolution, as we will discuss.

Organizational Growth vis-à-vis Open Data Ideals

At the time of this writing, OSM is ten years old and maturing rapidly. In its early days, OSM's focus was on simply gaining ground on the practicalities of collaborative spatial data production among a distributed volunteer base. Its members were the primary consumers of the map, so the relationship between creation and use was closely knit. Today OSM sits in a broader arena of actors where the implications of the ideals of open data come differently into view, requiring greater internally- and externally-coordinated work. Humanitarian activities have been a driver of OSM's evolution, in part because open data and participatory ideals align with humanitarian work, but also because disasters are catalysts for organizational innovation generally speaking [20]. Humanitarian events thus become useful lenses for isolating strategies that the wider OSM community has adopted to cope with growth in membership as well as geographic and thematic scope.

Disaster Events as Windows into OSM Evolution

The 2010 Haiti Earthquake was a significant event for OSM [27]. In this research, we take the collaborative work [24] happening on and through the map during this timeframe along with another major disaster—the 2013 Typhoon Yolanda—as windows into OSM's organizational evolution. To motivate rationale and foreshadow the discussion to come, we draw attention to the map data creation during the one-month periods following each event in Figure 1 (with the methods for how the visualizations were created delayed for narrative clarity). The orderings are remarkably different. Notice how in Port-au-Prince, the data by mappers—who are represented by color—are distributed across the area. In the Tacloban region of the Philippines, mapper activity is far more clustered, with little overlap. How this reordering came about was, on first glance, a straightforward technical solution, but we will show how it is rather a socio-technical outgrowth of OSM's tacit and explicit organizational objectives.

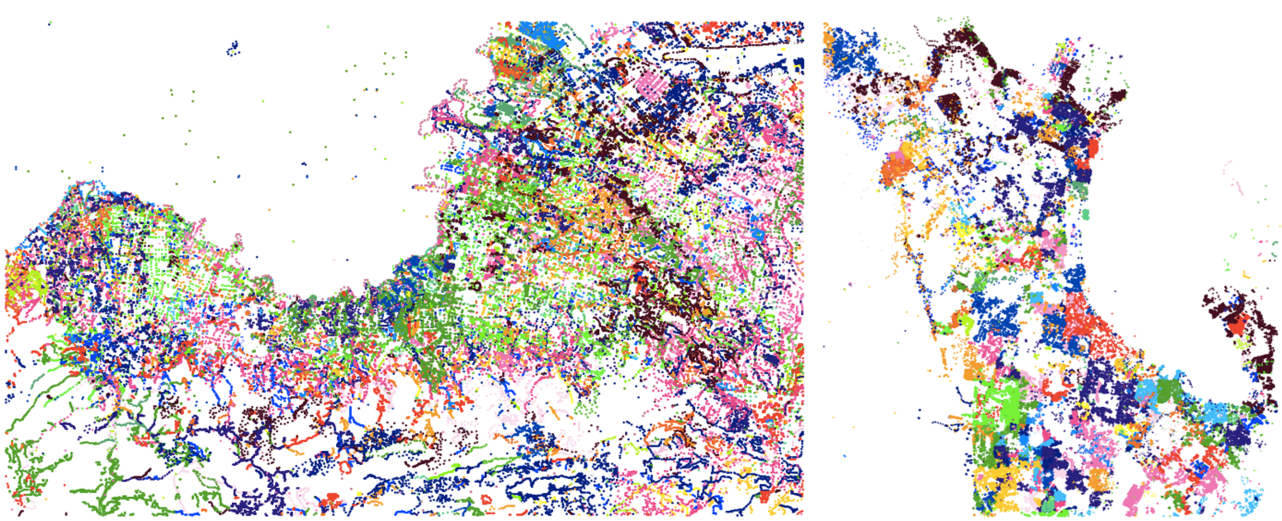

Figure 1. Marked differences in map data creation by users (colorized) across two disaster events reflect evolution in OSM organizing. Visualization compares 116k node additions in Port-au-Prince, Haiti (left) & Tacloban, Philippines (right).

Research Objective

Our aim in this paper is to capture the socio-technical evolution that these two "collaborative signatures" of map creation represent. We examine the production function of OSM during these two events to consider how and why it shifted over a four-year period.

RELATED LITERATURE

We review the OSM literature as it relates to organizational production functions. The Wikipedia literature gives insight into structuration processes that might have analogies for OSM. Finally, because we use the disaster frame to scope analysis, we discuss digital participation in such events.

OpenStreetMap

The literature describing the nature and work of OSM is relatively limited. Haklay [8] studied OSM data in comparison to official sources and found OSM coverage to be uneven but its quality generally high. As an organization, Mooney and Corcoran argued that OSM is "fragmented…comprising of individuals performing...tasks in isolation with little collaboration or interaction with each other" [21]. Though much of the work of editing OSM is indeed conducted by mappers working alone in front of their computers, we believe this simplified view ignores not only the manner in which mappers accumulate mapping work asynchronously over time, it also ignores the ways in which mappers articulate work "off the map." Lin [19] has emphasized the importance of OSM's State of the Map conference to share ideas and discuss issues. Hristova et al. [13] has looked at how in-person mapping parties promote sustained engagement. The published record of the range of ways that OSM mappers interact on and off the map is otherwise limited, but our analysis reveals that non-map forms of computer-mediated communication are important for establishing best practices, decide on metadata schemes, and otherwise organize as a community.

Wikipedia

In contrast, Wikipedia research has been extensive. As the community grew, governance policies became more complex and stabilized [6]. However, these processes for maintaining quality may now have deleterious impacts on Wikipedia's ability to retain new contributors. In particular, the growth of algorithmic monitoring and the increasing rigidity of policies have been shown to reduce the number of new contributors who continue editing [10]. Further, Wikipedia faces a significant challenge encouraging diversity among contributors. A 2011 study showed that only around 7% of editors were women and that unlike other online gender gaps, this did not seem to be improving [17]. In work related to our examination of OSM here, studies of editorial work during "high tempo" situations describe it as shared work by members with "profound differences in perspectives and norms" [16]. We will show that OSM has responded in its own adaptive ways to membership growth and diversity, and to the management of distributed responsibility across a loosely connected organization comprised of individuals with varying goals.

Digital Volunteerism during Disasters

Crisis informatics research has examined the emergence and convergence of "digital volunteers" who mobilize via social computing platforms to report on events, search and distribute situational information, and articulate unspoken needs on behalf of victims [1,28,29]. The work of digital volunteerism is diverse, but can be thought of as information tasks that can be done remotely; can resolve uncertainty through sleuthing; and can sometimes generate data by combining existing sources anew. Disasters often attract new digital volunteers; many disperse after an event is over, but several continue to help from event to event [29]. As we will see, OSM has drawn both experienced and new mappers to humanitarian causes.

THE STUDY

We now describe the disaster events that serve as the temporal and spatial constraints for examining OSM's evolution. We describe data collection and provide an overview of the OSM data structure.

Two Disaster Events: Overview of Conditions

Haiti Earthquake

Just before 5pm local time on January 12, 2010, a 7.0 magnitude earthquake struck Haiti. The epicenter was 20 miles outside of the capital Port-au-Prince. The heavy block and concrete style construction of the capital city—intended to withstand hurricanes—collapsed and caused massive loss of life and injury. It is now estimated that over 40,000 people died and over 1 million were displaced. As many as 40% of Haiti's civil servants were injured or killed, and the majority of government buildings were damaged or destroyed [14], rendering access to government documents (including maps) difficult. The UN along with other international organizations launched one of the largest disaster relief efforts in history.

Typhoon Yolanda

On November 8, 2013, one of the strongest storms ever recorded hit the Visayas region of central Philippines: Typhoon Yolanda (also known as Haiyan). The storm itself affected more than 8 million people and killed over 7,000 [31]. The subsequent flooding, landslides and storm surge affected an estimated 14 million people, over 14% of the country's population [30]. Over 50 international groups have contributed to the response and recovery efforts [30], which, at the time of this writing, are ongoing.

Data Collection

Our methods rest on significant participant observation in OSM between 2009 and 2014, and content analysis of blogposts, listserv conversations, and wiki pages. We also conducted quantitative analysis of records in the OSM database during the one-month long periods following the onset of each disaster. The dataset includes data records of all edits from January 12-February 12, 2010 for Haiti, and November 8-December 8, 2013 for the Philippines.

To analyze cartographic activity, all of Haiti's territory was included, as OSM mapping covered the country. In the Philippines, mapping and therefore analyses focused on the Typhoon-affected areas (Figure 2). Data were imported into an instance of the Mongo document database, where we used Ruby, R, Splunk, and custom queries to analyze.

Figure 2. Bounding boxes for Haiti (left) & Philippines (right)

OSM Data Structure

We analyzed all unique versions of *nodes*, the building blocks of all OSM data. A node represents a single geographic point. Roads, paths, buildings, and areas are represented as a string of several nodes; this relationship is called a *way*. Our database contains entries for every version of a node, whose structure is shown below (Figure 3). This allows us to reconstruct the modification history of objects. The *changeset id* for each node is unique to a set of edits committed to the database by a user. Large automated data imports that represented macro features like country boundaries were not included because they did not represent contributions by individual mappers and therefore skewed results.


```
{"id": 1333723135,
"lat": 18.5342186, "lon": -72.3673319,
"timestamp": 2013-06-24T02:49:56Z, "changeset":
16677898,
"version": 2,
"user": "jaakkoh_massedits",
"tags": {"source": "cosmha_iom",
          "amenity": "drinking_water",
          "pump_type": "manual",
          "operational_status": "open",
          "water_quality_issue": "yes"} }
```


Figure 3. A distilled JSON representation of an OSM node

ANALYTICAL DESCRIPTION

We present details of OSM response first to the Haiti event, including reports of numbers of participants and edits to the database, as well as descriptions of the nature of their mapping. We discuss how collaboration "off the map" happened. In the 3.5 years between Haiti and Yolanda, several changes took place within OSM, some in direct response to the Haiti experience. We then report on activity in the Philippines and how behaviors changed.

OSM in Response to the Haiti Earthquake

The Haiti Mappers

In the month following the Haiti earthquake, 491 OSM users contributed 12,858 distinct changesets. This included edits to 1,217,448 nodes, 956,725 of which were new additions (Figure 4). The mappers created the most detailed map of the quake-affected area in existence [27]. User contributions varied dramatically, with a mean of 2,471 nodes edited, but a median of just 223, exhibiting a long-tail distribution.

Of the total Haiti mappers, 24% were new to OSM, seemingly drawn by a desire to help that we see in other digital disaster response efforts [28]. New mappers—those who joined OSM following the disaster—contributed 10.4% of all node edits, meaning that experienced mappers—those who already had accounts at the time of the earthquake—conducted the vast majority of mapping.

OSM Experience Level	Number of Mappers	Total Nodes Edited	Total Changesets Made
New	119	126,392	891
Experienced	372	1,091,056	11,967

Table 1. Experience & Contribution during the Haiti window

"High Tempo" Mapping: Collisions in Simultaneous Work

Though their accomplishments were remarkable, the high tempo activity by numerous mappers in a constrained geography was problematic. Moments of conflict and collision happen at a very low-level, in pieces and parts of a

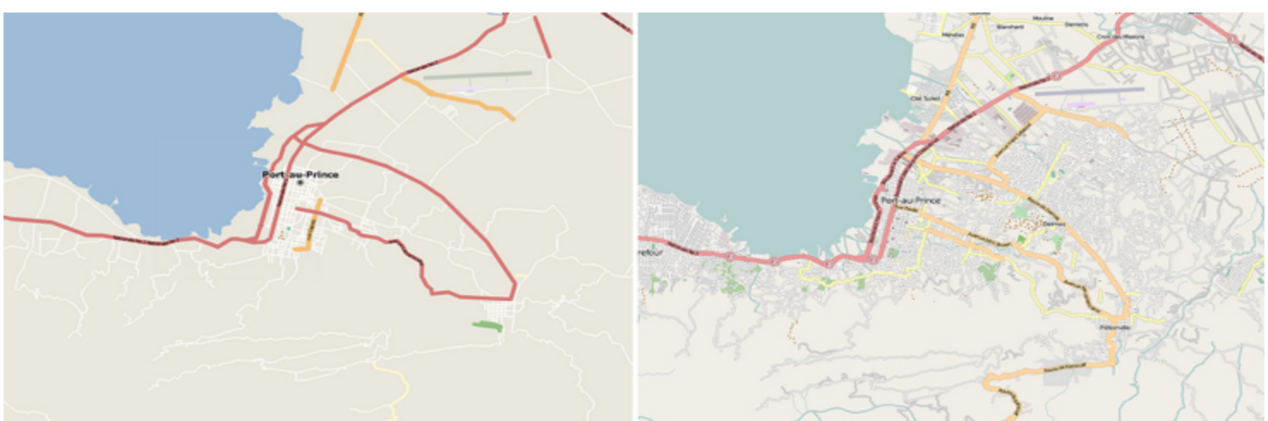

Figure 4. Rendering of Port-au-Prince before & after earthquake, a 9-fold increase in *node* content. 1/9/2010, 1/20/2010 [15]

building or road, which can be hard to detect even by the mappers. We illustrate these moments of congestion to convey the nature of high tempo work in OSM.

Figure 5 shows a simple case of duplication that eventually reached resolution. Here, two changesets less than 1 minute apart contained 4 duplicate road sections created by two mappers (one red, one blue). The duplicate sections in this case were resolved when the second mapper returned to delete the first mapper's traces 4 hours later.

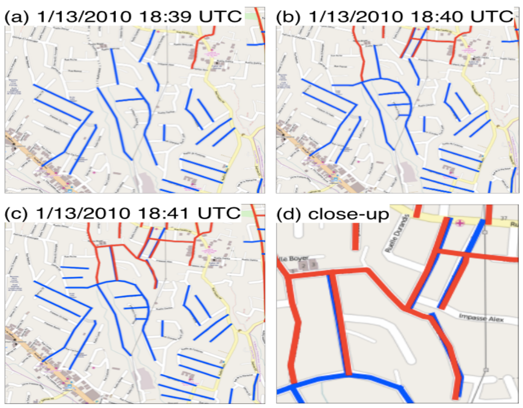

Figure 5. Four overlapping roads by two mappers

A second case (Figure 6) shows how a road was produced by three mappers, but the entirety of the tracing remained unresolved. Here, a remote mapper working the southwest area of Port-au-Prince mapped and tagged the highlighted feature as "highway: unclassified" (Fig 6b). Three minutes later, a different mapper created another way for the same feature, but tagged it "highway: road" (Fig 6c). Working at nearly the same time, the mappers probably were unaware of the collision. Five hours later, a third mapper deleted the road by the second mapper, and created yet a third road that used only some of the nodes from the first road. This final road was also tagged as "highway: unclassified," meaning that the resolution of the duplication remained incomplete.

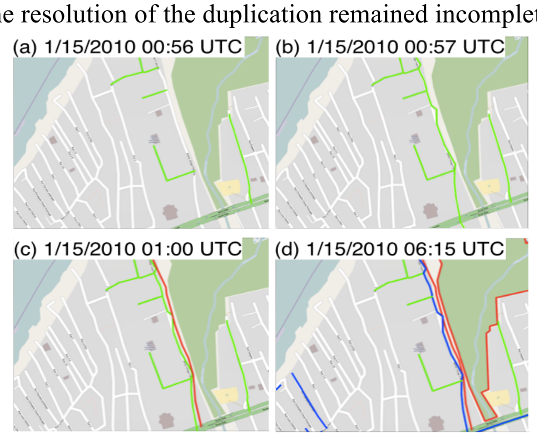

Figure 6. Overlapping roads created by three mappers

These selections illustrate how high-tempo mapping work in constrained geographies can result in map collisions. Though collisions often get resolved, the experiences reported by mappers who worked on the Haiti event in combination with evidence of duplication throughout the database indicate that valuable time was spent on producing the map erratically. These collisions were not a matter of inexperience: practiced mappers were at work here. What is important about collision and duplication is that maps for crisis response are being created for immediate use on-the-ground: maintaining map integrity as its features are being rapidly produced by evermore participants takes on new levels of importance and challenge.

Collaboration "Off the Map"

The Haiti earthquake mappers loosely organized their work through public listservs, IRC discussion, and a wiki page—what we call collaboration "off the map." There they strove to motivate the community to contribute with prioritization of features that international responders especially needed, including spontaneous settlements of those displaced by the earthquake ("tent cities"). They also needed to manage the high volumes of donated satellite imagery that was the basis for their mapping. Though those tasks were critical to the overall effort, they still could not address the pervasive problem of "in-the-map" collisions. Indeed, one of the organizers wrote to the general OSM mailing list asking:

> What tools are available to see in real time which areas have been mapped recently? … Any ideas on another tool that would allow people to "check out" an area for a period of time? … This would be helpful for general coordination among us mapping in Haiti.

This request was unfulfilled, though not for long, and the challenges that OSM faced as it found itself in new territory were coming into view.

Changes to OpenStreetMap in the Wake of Haiti

Lessons from Haiti gave rise to socio-technical changes within the organization. In addition, other changes were afoot that, in total, paint a picture of evolution toward more clearly articulated implications of OSM's ideals.

The Formalization of the Humanitarian OSM Team (HOT)

OSM was the de facto basemap for humanitarian groups working in response to the earthquake. It was used by first responders during search and rescue, to plan and manage camps for displaced Haitians, and in response to the Cholera outbreak that began in October 2010 [27]. A group that called themselves the Humanitarian OpenStreetMap Team (HOT) traveled to Haiti six times with long stays during the 18 months following the disaster to help not only responding organizations adopt OSM, but also to generate Haitian interest in developing a local OSM chapter. They hoped that a local community could maintain and update the map for the indefinite future. During this time HOT worked closely with partners in humanitarian organizations to understand their data needs and adapted OSM's practices accordingly, including building a new data model of map features as well as releasing the data more rapidly [27]. This objective represented an important turn, as it shows how OSM's purpose was expanding from a primary focus on map creation to now include a focus on situated, localized data use by organizations that, in this case, a newly instituted Haitian OSM community could support.

HOT formally registered as a non-profit organization that following August, just 7 months after the earthquake [27]. It now employs an executive director, consultants, and a Board of Directors elected by the OSM membership. It has worked in a number of countries in Africa and Asia where, as in Haiti, they sought to encourage the growth of local OSM communities and bring contributors to the effort. By the time of Hurricane Yolanda, HOT had become the locus of activity within the OSM community for disaster response and international development concerns.

Innovating to Manage High Tempo Mapping Conflict

The map congestion experienced in the aftermath of the Haiti earthquake led the newly formed HOT to create *The OSM Tasking Manager* to help mappers more efficiently coordinate simultaneous work. This innovation coincided with the growing popularity of microtasking as a solution to conduct distributed work (e.g., [1]). The design resembles microtasking environments that have people working in parallel with little knowledge of what others are doing. Administrators create "jobs" for a large geographic area, with written instructions about the features to be prioritized and tagging schemas if they differ from global OSM practice. The Tasking Manager then divides each job into a grid of "tasks" which mappers can select (Figure 7). Yellow means the task is taken; red denotes that it is complete but awaits "validation" by another; green denotes that it has been inspected in the OSM database for completeness and compliance.

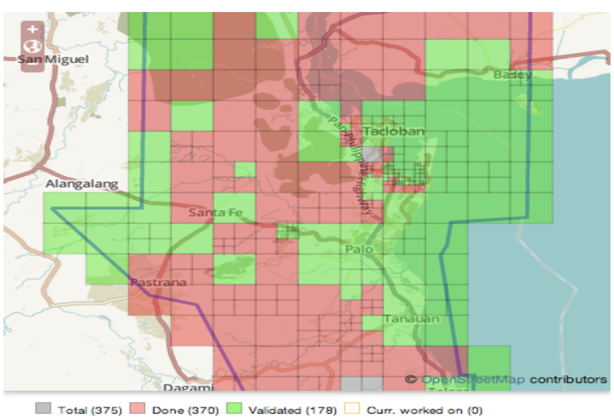

Figure 7. The user interface of the Tasking Manager [12]

The first version of the Tasking Manager launched in September 2011 and was used to help in a number of smaller emergencies, but was not used during a crisis on the scale of the Haiti earthquake until Yolanda struck.

Other Changes within OSM at Large

The period between the disasters in Haiti and Yolanda was also an important time of change for the wider OSM community. The growth of the organization involved in contributing to the map and the users who relied on its data were driving a number of parallel efforts to meet desires for attracting and retaining new participants, and making data more reliable and accessible.

Legal Status of OpenStreetMap Data. Underway for several years but finalized in September 2012, OSM moved away from the Creative Commons and towards ODBL—the Open Database License. This was due to concern that Creative Commons Licenses were created for works of art rather than databases and therefore did not properly protect the data and those who created it. This indicates further orientation to downstream activities of data consumption, and an articulation about how operations can work in an open data goal-driven environment.

Broadening Appeal. In May 2013, an updated OSM editing interface for mapping was released to make it easier for new participants to begin mapping. Improvement of the openstreetmap.org website was intended to achieve the same goal. Both efforts were conducted by Mapbox with a grant from Knight Funding, whose announcement reveals the growing support and stronger orientation toward not only data creation, but also consumption:

> Our goal is to use this investment to make it easier to add data to OpenStreetMap, make OpenStreetMap.org more social to support the community as it continues its rapid growth, and make it easier for people to get data out of OpenStreetMap to make their own maps.

A website called LearnOSM was built by HOT to consolidate the great deal of documentation and training materials to help new mappers. Launched in June 2012, LearnOSM materials are translated into 9 languages to reach people in developing regions. In another attempt to broaden appeal, a new email list was launched in June 2013 after discussion at OSM conferences earlier that year to promote inclusion of a more diverse mapping population, particularly women. Some estimates put the number of female contributors to OSM as low as 3-5% [24], below even Wikipedia's numbers. These efforts again suggest an orientation toward better mapping contributions by a diverse population, a major driver for sustaining a viable community that would hopefully not suffer from some of the problems Wikipedia faces in its exclusive user base.

Other Support for Collaboration. In its own parallel effort distinct from the Tasking Manager, the development of *Notes*, a drop-pin annotation feature, is described as an issue-tracking feature to point out improper tagging or to suggest additional information that could be added to the map. In addition, guest users are able to create Notes, such that "everyone can add their local knowledge to the map" without having to learn how to edit in OSM. The Notes feature enables meta-conversation that sits on top of the map and is not encoded in the database. The creation of this new feature suggests that OSM is exploring new approaches to collaboration that are based on the map but take place a step removed from the database itself. We will return to the use of Notes in our discussion of Typhoon Yolanda.

Resistance to Expansion of Governance Structures. Finally, we note where changes did not take place. Unlike Wikipedia, which has sought to manage community growth through the creation and clearer articulation of policies [2], the OSM governance structure remained largely the same during this period and still today. The mission and responsibilities of the OSM Foundation, a small elected group and its working groups, has remained narrowly scoped: to maintain server infrastructure, host the annual State of the Map conference, and act as a legal and fiscal entity for the effort. Debates on the OSM listserv and at the State of the Map indicate a strong desire to avoid what are seen as bureaucratic solutions to project governance in favor of social and technical approaches.

OSM in Response to Typhoon Yolanda

Almost four years after the Haiti earthquake, Typhoon Yolanda struck the Philippines. Over this period, OSM membership quadrupled to over 1.2 million users. The response to the event was significant with the work of the volunteer mappers led by the now well-established HOT.

The Yolanda Mappers

In the month following the typhoon, 1,574 mappers made 3,648,537 node edits, including adding 3,294,981 new nodes to the map, averaging 2,292 edits per mapper, with a median of 426. This again signals a long-tail distribution. Figure 9 shows the results of their mapping contributions.

OSM Experience Level	Number of Mappers	Total Nodes Edited	Total Changesets Made
New	663	491,404	14,562
Experienced	911	3,157,143	22,266

Table 2. Experience & Contribution during Yolanda window

We see a roughly 3-fold increase in both the number of mappers as well as node edit contribution between Haiti and Yolanda. This increase tracks with the overall OSM membership growth and was likely also influenced by the efforts of HOT. A comparison of the rapid mobilization of remote mappers and the work they conducted is depicted in Figure 8. Examining both disasters at once reveals a dramatic spike in activity that follows hazard onset. Mapping in the Philippines begins before onset because typhoons yield advance warning. The number of daily users editing follows this same trend, peaking with 85 users and 28 users/hour in the Philippines and Haiti, respectively.

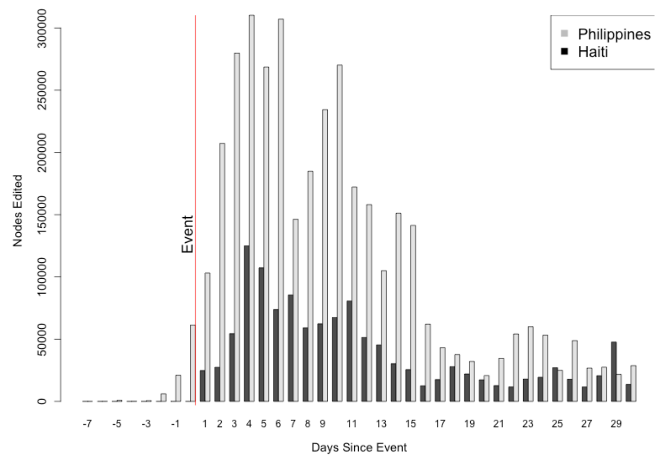

Figure 8. Number of nodes/day for both events

As with Haiti, we see that during Yolanda, experienced users made the majority of contributions with new mappers contributing just 13.5% of the total node edits.

Use of the Tasking Manager During Yolanda Response

HOT coordinated the OSM community's response during Yolanda using email and an IRC chatroom to motivate volunteers, discuss imagery issues or finer technical details of mapping, and direct contributors to the Tasking Manager. Because the OSM database does not record whether a mapper's edits are directly related to the Tasking Manager, we compared the history of the Tasking Manager database with the OSM changeset history to determine how much this new tool was used. During the study window, 1,354 or 86% of the mappers checked out tasks and made

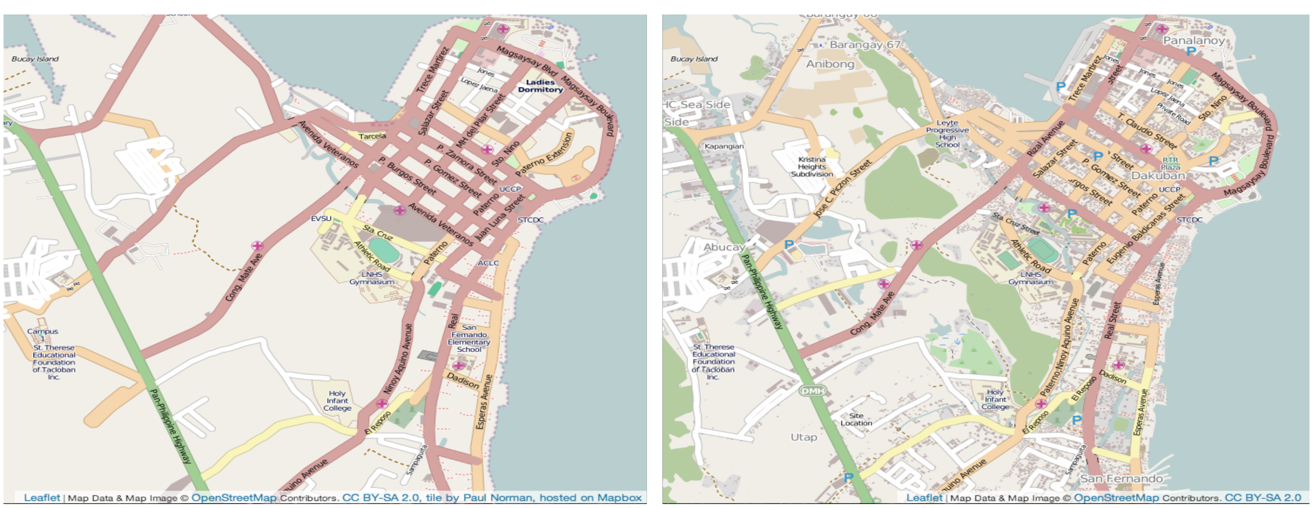

Figure 9. Rendering of Tacloban before & after Yolanda, a 11.5-fold increase in *node* content. 11/7/2013,11/ 17/2013 [22]

changesets that geographically coincided with one or more of the Yolanda Tasking Manager jobs. Mappers who did not appear in the Tasking Manager database contributed only 6% of the Philippines OSM changesets. These results indicate that the Tasking Manager was widely used by those who participated in remote mapping.

A changeset can illustrate not just what mappers edited but *how* they edit the map. We employed several metrics to study editing patterns, including nodes per changeset, area of the changeset bounding box, and the quotient of the two: what we call the "density of the changeset." Within both datasets, some changesets reflect large automated imports of data. For the purposes of the node count, area, and density analyses, the changesets we observed were greater than $1m^2$ and limited to those containing less than 10,000 nodes to avoid inclusion of automated imports.

The median area for changesets in the Philippines was $0.5km^2$, significantly lower than for Haiti where the area was $3km^2$. In addition, changesets in the Philippines were, on average 9 times denser than changesets in Haiti. These numbers indicate that the Philippines mappers tended to be more geographically focused, a likely impact of Tasking Manager use. Furthermore, when looking at simply the number of changesets that geographically overlap, we see that in the Philippines, an average changeset intersects with 67% fewer changesets than in Haiti.

Returning to the first visualization offered in this paper (Figure 1), we see the vivid before-and-after effect of the Tasking Manager intervention that illuminates these figures reported here. (The bounds of each figure were chosen so that they would have a common number of nodes—116K—for consistent comparison.) For Haiti, there is a great deal of dispersed mapping. In the Philippines, however, a different ordering emerges, with patterns of concentrated mapping by people working (virtually) side-by-side.

Lessons Learned from the Tasking Manager in Yolanda

HOT reflected on the Yolanda experience in early 2014 (not long before the time of this writing), and has developed a set of priorities to improve the Tasking Manager. Among them include a new addition by which mappers working on the same task can talk with each other—asynchronously through comments—to discuss uncertainties. Identification of this priority is consistent with something we anticipate: Our observation is that the microtasking environment so reduced mapping collisions that opportunities for deliberation and organizational learning from those erstwhile "entanglements" are lost. In other words, when seeking to address collisions, mappers may be required to revisit mapping activity conducted by themselves or others; reflect on available information about the geographic area and make decisions about how best to represent it in the map; or work with other members of OSM to resolve disagreements. The task comments could be a way to reintroduce benefits of collaboration without cost to the efficiency that was otherwise gained for time-critical mapping.

Also on the list of priorities was an item that we think signals a very important concern about higher order issues of coordination. Since all the jobs are put into one instance of the Tasking Manager no matter where on the globe the work is, there could be yet a new form of traffic within the confines of the microtasking tool. Here, the team is considering how to specify division for bigger geographic or population segments ("for Haiti," or "by the ugrad course at University X") to mitigate this potential problem.

These solutions reinforce three things with respect to analysis of OSM's evolution: 1) that OSM is thinking about how to scale as its mapping population grows; 2) that the database, though powerful for what it represents as an open spatial data repository, will become only increasingly problematic with respect to how it supports and reveals social interaction as OSM grows; and 3) that OSM is struggling again with technical solutions for collaboration.

Use of the Notes Feature During Typhoon Yolanda

Recall that the Notes feature was invented in April 2013 as yet another solution for supporting collaborative activity. There was some use of Notes in Yolanda, though not extensive (30 registered accounts used it compared to the 1554 total mappers for that event). There were 241 Note instances during the Yolanda window.

Some registered users were striving for better coordination through Notes to resolve uncertainties in the aerial photography or unknowns on-the-ground. For example, here we see an interaction about language translation from Tagalog to English for a missing bridge name. A guest who is perhaps Filipino notes a missing feature; two OSMers move the problem to solution in two steps, and resolve it:

Anonymous (2013-11-15 21:30:39) [action:opened]: mao ni ang buray bridge

GOwin (2013-11-16 05:45:14) [action:commented]: translation: "this is Buray bridge"

ianlopez1115 (2013-11-19 05:55:57) [action:closed]: bridge name added

In other Notes exchanges, OSM mappers tried to encourage new membership. However, some guests demonstrated that they see OSM *as a map to appropriate and personalize*, not a map to directly create. They are acting more as "involved viewers" rather than data-producing mappers who view the map more functionally. One tagged his residence "Home Sweet Home," while another offers map data combined with marketing:

BILAR BULIK (2013-11-28 06:06:59):
created by BILAR BULIK at Nov 28, 2013 06:02 AM
BILAR COCKPIT ARENA
OWNED and MANAGE BY:
CAPTAIN ELMER BROXX NAMOCAT
THE Finest STALKER FARM
BREEDERS OF BILAR BULIK ENGLISH DOME
HOME OF BILGBA
BILAR GAMEFOWL BREEDER ASSOCIATION…

Perhaps this new user would compel Geertz [7] to take up OSM as a site of study: Bilar Bulik breeds cocks and runs a cockfight arena. This is the only Note he adds, and though registered as an OSM user, he makes no edits to the database. It would seem that he just wants to make his mark on the world—that is, the OSM map as world.

Finally here we see a guest trying out the Note feature by typing in a QWERTY patterned non-sensible entry. `GOwin` sees this as a chance to explain his imagined use of Notes:

> **Anonymous** (2013-12-05 01:56:40) [action:opened]: asdasdasdasd
>
> **GOwin** (2013-12-05 03:01:24) [action:closed]: Use this feature to report an error in the data or to give some additional information, for instance the name of a street or an address etc. When available, an OSM contributor will attempt to resolve it.

Though it was clear that some mappers strove to make Notes a mechanism for coordination, norms around its use have yet to stabilize. We predict that Notes will continue to have a degree of volatility because guests might see Notes as an opportunity for personalization, or for noting information that is transient, rather than permanent (such as a school marked as a temporary disaster shelter). This differs from the prevailing view of the map as encoding physical structures that are permanent and visually verifiable. We see both the conceptualized and actual use of Notes as expressions of a desire from multiple directions to support legitimate peripheral participation [18].

CONCLUSION

OpenStreetMap, like Wikipedia around 2007, has been undergoing rapid growth and has had to find ways to both cope with and maximize this opportunity. In the early days of OSM, many of the important users of the data were members of the community and thus also directly involved in the creation of the data. As new groups began to rely on OSM data, such as the humanitarian organizations operating in Haiti and the Philippines, the organization has had to look for ways to make itself more accessible to outsiders. OSM has attempted to address these twin challenges of community growth and accessibility to an array of new users (both mappers and consumers) by focusing on the usability of its tools, addressing legal questions around usage and distribution of the data, and working to attract and retain new participants.

We find evidence that OSM is increasingly orienting to the obligations that come about when open data is an organizational value. We argue that to see such a value through, an organization must come to see itself as part of a constellation of data use relationships that place demands on such issues as data integrity, availability and accessibility. Unlike commercial efforts that are perhaps necessarily oriented toward its customers, volunteer-based open data efforts may be able to suspend orientation to that goal while they are bootstrapping data production and building their community, but to remain sustainable and indeed to have the meaning that open data values connote, a hard turn toward the demands of data consumption is inevitable. As a participant in the State of Map Conference 2010 said of OSM to great applause:

> Aren't we just hitting puberty? We were born 4 years ago; we are growing up…Look at what we've accomplished. We are growing appendages and we don't really know what they do. And we are trying to figure them out and occasionally they don't work like they should…[but] we've put people in Haiti, we're putting people in Kibera…[so] we're growing up…
>
> [But] it's not just [about] the language *we* are speaking, it is about the language *that everyone else is speaking*. So yeah, we do need to have that culture where we just don't "do it," but we also need to present that to the [for example] "assistant under secretary of management" for the UN, because we do have to speak that language as well.

To these ends, OSM has made an effort to first reorient in practical terms to get its collaborative work in hand. Different parts of the organization have created different tools to manage work in an effort to "realign" [3] with its expanding user base as the organization envisions how to proceed in relation to larger goals of data reuse that are coming into view. In effect it is creating and maintaining the kind of cyberinfrastructure that Bietz, Ferro and Lee describe [3]. However, we speculate that some of OSM's current solutions might be closing out opportunities that spring from *beneficial entanglements*. Some entanglement in distributed work can benefit organizational evolution, because it is on those virtual street corners [33] of the map where action that later supports substantive deliberations about mapping practices, policies and "speaking other languages" of all stakeholders are first staged.

However, the work of OSM is hard to assess quantitatively or qualitatively because so much of it is written in at the lowest levels *into* the database. Even core members struggle with understanding the nature of the work and how it comes about because it is simply so hard to "see" it. The power of the database as the repository for open spatial data has both propelled OSM forward, but also creates special challenges. OSM as an organization cannot easily separate itself from how it produces and encodes its data.

To move forward then, OSM's solutions are sitting "on top" of the map. Though this abstracted articulation work thereby is more visible, and easier both for participants to conduct and researchers to observe, the view of the work at the level of map genesis is still obscured. This paper, en route to explaining organizational evolution, offers some measures for mappers and observers to understand map-based interaction as it is inscribed in the database, and we hope to further facilitate analyses of OSM's rich organizational life. We see such contributions as belonging to the important call to better understand matters of geospatial data [11]—its creation, its reuse, and the social infrastructure that surrounds it—as a growing issue in human-computer interaction.

REFERENCES

1. Bernstein, M., Little, G., Miller, R., Hartmann, B., Ackerman, M., Karger, D., Crowell, D., Panovich, K. Soylent: a word processor with a crowd inside. *In Proc. of UIST* (2010), 313-322.
2. Beschastnikh, I., Kriplean, T., McDonald, D. W. Wikipedian Self-Governance in Action: Motivating the Policy Lens. In Proc of ICWSM. (2008).
3. Bietz, M., Ferro, T., Lee, C. Sustaining the development of cyberinfrastructure: an organization adapting to change. In Proc of CSCW (2012), 901-910.
4. Budhathoki, N., Haythornthwaite, C. Motivation for open collaboration crowd and community models and the case of OpenStreetMap. American Behavioral Scientist 54, 5 (2013), 548-575.
5. Chilton, S. Crowdsourcing is radically changing the geodata landscape: case study of OpenStreetMap. In Proc. of UK Cartographic Conf, (2009).
6. Forte, A., Bruckman, A. Scaling consensus: Increasing decentralization in Wikipedia governance. In Proc. of HICSS, 2008.
7. Geertz, C. Notes on the Balinese Cockfight, Chapter in the Interpretation of Cultures. Basic Books (1973).
8. Haklay, M. How good is volunteered geographical information? A comparative study of OpenStreetMap & Ordnance Survey datasets. Environment & Planning B: Planning & Design 34, 4 (2010), 682-703.
9. Haklay, M., Antoniou, V., Basiouka, S., R Soden, R., Mooney, P. Crowdsourced Geographic Information Use in Government. World Bank (2014).
10. Halfaker, A., Geiger, R. S., Morgan, J. T., Riedl, J. The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline. American Behavioral Scientist. (2012).
11. Hecht, B., Schöning, J., Erickson, T., Preidhorsky, R. Geographic Human Interaction. Extended Proc, of CHI (2011).
12. HOT Tasking Manager. Typhoon Haiyan. http://tasks.hotosm.org/job/350.
13. Hristova, D., Quattrone, G., Mashhadi, A. J., Capra, L. The Life of the Party: Impact of Social Mapping in OpenStreetMap. ICWSM (2013).
14. Human Rights Watch (HRW). World Report 2011. Seven Stories Press (2011).
15. ITO World Blog. January 20, 2010. http://itoworld.blogspot.com/2010/01/mapping-crisis-openstreetmap-response.html.
16. Keegan, B, Gergle, D, Contractor, N. Hot off the wiki: dynamics, practices, and structures in Wikipedia's coverage of the Tōhoku catastrophes. In Proc. of WikiSym, (2011).
17. Lam, STK., Uduwage, A., Dong, Z., Sen, S., Musicant, D. R., Terveen, L., Riedl, J. WP: clubhouse?: An exploration of Wikipedia's gender imbalance. Symposium on Wikis & Open Collaboration. (2011).
18. Lave, J., Wenger, E. Situated Learning: Legitimate Peripheral Participation. Cambridge Univ, (1991), 138.
19. Lin, Y. A qualitative enquiry into OpenStreetMap making. New Review of Hypermedia and Multimedia 17, 1 (2011), 53-71.
20. Mileti, D.S. 1999. Disasters by Design: A Reassessment of Natural Hazards in the United States. Joseph Henry Press. Washington D.C.
21. Mooney, P, & Corcoran, P. How social is OSM. In AGILE Conf on GIS. 2012. (24-27).
22. OpenStreetMap Development. http://pierzen.dev.openstreetmap.org/hot/leaflet/OSM-Compare-before-after-philippines.html.
23. Priedhorsky, R., Jordan, B., Terveen, L. How a personalized geowiki can help bicyclists share information more effectively. In Proc. of WikiSym, (2007), 93-98.
24. Schmidt, K., Bannon, L. Constructing CSCW: The First Quarter Century. Journal of Computer Supported Cooperative Work, (2013), 22 (4-6), p. 345-372.
25. Schmidt, M., Klettner, S. Gender and Experience-Related Motivators for Contributing to OSM. In Proc of ACTIVITY (2013).
26. Soden, R., Budhathoki, N., Palen, L. Resilience-Building and the Crisis Informatics Agenda: Lessons Learned from Open Cities Kathmandu. In Proc. of ISCRAM, (2014).
27. Soden R., Palen, L. From crowdsourcing to community mapping: the post-earthquake work of OpenStreetMap Haiti. In Proc. of COOP, (2014).
28. Starbird, K. Palen, L. "Voluntweeters": self-organizing by digital volunteers in times of crisis. In Proc. of CHI, (2011), 1071-1080.
29. Starbird,K, Palen,L. Working & Sustaining the Virtual "Disaster Desk." Proc. of CSCW (2013), 491-502.
30. United Nations Office for Coordination of Humanitarian Affairs (UNOCHA) 2013. https://docs.unocha.org/sites/dms/CAP/SRP_2013-2014_Philippines_Typhoon_Haiyan.pdf
31. USAID Yolanda Factsheet - http://www.usaid.gov/haiyan/fy14/fs20
32. White, J., Palen, L., Anderson, K. M. Digital mobilization in disaster response: the work & self-organization of on-line pet advocates in response to hurricane sandy. In Proc. of CSCW, (2014), 866-876.
33. Whyte, W. H. The Social Life of Small Urban Spaces. Chicago. (1980).