Datafying Disaster: Institutional Framings of Data Production following Superstorm Sandy

Abstract

In the wake of disasters, communities organize to produce spatial data capturing knowledge about the disaster, and to fill gaps left by formal emergency responders. The ways communities impact overall response efforts can produce inequalities, disempowerment, or further marginalization. Increasingly, this organizing and knowledge production occurs through digital technologies, and recently “digital humanitarianism” has become an important suite of such technologies. Digital humanitarianism is comprised of technologies like the crowdsourced crisis mapping platform Ushahidi, and the community of volunteers Humanitarian OpenStreetMap Team which focuses on the amateur-generated global basemap OpenStreetMap. Digital humanitarianism is shifting the ways needs and knowledges are captured and represented as data following disasters. These transformations raise important questions for geographers interested in the socio-political and institutional processes that frame data production and representation.

In this article, I contribute to geographers’ efforts to understand the institutional and community-based politics that frame the types of data that are produced in disaster contexts. I do so by drawing on an ethnographic project that took place in both Washington, DC and New York City after Superstorm Sandy in 2012. I show that digital humanitarians produced data in the Rockaway Peninsula of New York in response to perceived gaps on the part of formal emergency responders. In so doing, they represented needs, individuals, and communities in ways that local community advocacy organizations found problematic. These findings shed light on the politics and struggles around why particular datasets were produced, and the motives behind capturing particular disaster-related needs and knowledge as data.
Introduction

Following large, disruptive events such as disasters, communities often self-organize to fill gaps in the responses of formal actors, such as through sharing resources and providing personal assistance (Stallings and Quarantelli 1985; Birch and Wachter 2006). This form of community empowerment, which has long been a topic of research (Wachtendorf and Kendra 2006), can result in the strengthening of interpersonal ties and the establishment of socio-spatial networks that can be politically powerful. Such organizing efforts increasingly occur through the use of digital technologies such as social media and Big Data (Mayer-Schönberger and Cukier 2013), text messages on mobile phones, crisis mapping, and crowdsourcing platforms. Despite their increasing prevalence, these technologies, collectively referred to as “digital humanitarianism” (Burns 2014, 2015; Meier 2015), have mixed and uneven impacts on community organizing in disaster contexts (Brandusescu, Sieber, and Jochems 2015; Read, Taithe, and Mac Ginty 2016). Since disasters are important moments for social and institutional upheaval and the re-establishment of new norms and relations, it is important to consider the institutional and community-based forces that frame the ways in which people use digital technologies to enact recovery processes. The indeterminate and often contradictory impacts of digital humanitarian technologies can deepen inequalities, disempower, or further marginalize, and this complexity reflects digital technologies more broadly (boyd and Crawford 2012; Straubhaar et al. 2013; Graham 2014). Research on critical GIS and the geoweb has elucidated the forces that set limits around what is seen as a possible or “legitimate” use of technology, data, and maps (Elwood 2006; Johnson and Sieber 2011), yet these considerations have to date had limited impact on digital humanitarian and Big Data research (Crampton et al. 2013).

In this article, I contribute to research underscoring the institutional and social forces and processes that frame the types of data and representations used to affect social change. In so doing, I build on geographers’ efforts to understand the ways that communities use digital technologies toward achieving socio-political goals, such as addressing social injustices. I report on a research project conducted 2012-2013 which observed the ways digital and physical community organizations in New York City engaged digital humanitarian technologies to fill gaps in the formal response to Superstorm Sandy; most importantly, these complex relations cultivated tensions between digital humanitarians and the two other parties of community organizations and the formal responders. I argue that the contestations around Sandy-related needs and knowledge representation led to very different data being produced by different stakeholders. In fact, one cannot understand the data that were produced – and the specific languages and descriptions that were employed – without attention to these struggles around knowledge and need representation.

I begin by situating the current article within research on digital labor in disasters. I pay particular attention to research showing the socio-spatially variegated implications of digital technologies, especially around the ability to effectively leverage technology for social change. These inequalities foreground digital technologies as a key site and means for struggles over social change and the decisions that impact communities. I follow this by describing the methodology and empirical context of this project. With institutional and community-based framings of data in mind, I make two substantive

---

1 I use the term “digital labor” here to invoke recent Marxist readings of crowdsourcing, attentional economies, and the commodification of web-based activities (Scholz 2012; Fuchs and Sevignani 2013; Terranova 2014), all of which underpin the processes at work in this article.
arguments related to digital humanitarian technology use in the response to Superstorm Sandy. The first is that the institutional frameworks and established work flows of formal responders elicited new data practices by community organizations, largely conducted through digital humanitarian technologies. Second, other organizations resisted these new data practices and the ways digital humanitarians captured and represented their needs and knowledge. In particular, the proprietary nature of much of this work – specifically, the practice of putting data behind paywalls – disenfranchised communities on the Rockaway Peninsula. This led some individuals to contest digital humanitarian approaches to data collection and representation. Lastly, I conclude by briefly enumerating two recommendations for how policymakers and digital humanitarians may rethink their engagements with such technologies.

**Digital Labor in Disasters**

In contrast with early boosterist claims that new digital humanitarian technologies promote democracy, liberation, and empowerment (e.g., Meier and Munro 2010; Zook et al. 2010; Crowley and Chan 2011), recent research has begun to re-conceptualize digital humanitarianism as an uneven, contested, and sometimes problematic socio-political development. Humanitarian data production and interpretation require specialized and situated knowledge and skills that are not only inaccessible to many digital humanitarian technologists (Bhroin 2015; Finn and Oreglia 2016), but they also enroll broader data relations such as gender discrepancies (Stephens 2013; Cupples 2015) and political-economic imperatives (Burns forthcoming; Thatcher, O’Sullivan, and Mahmoudi 2016). Many have shown that digital humanitarianism impacts response efforts in unintended ways (Currion 2010; Jacobsen 2015) and can potentially expose vulnerable populations to increased risk (Shanley et al. 2013; Haworth and Bruce 2015; Raymond et al. 2016). Humanitarian data are produced in institutionally-specific ways that mirror political imperatives, sometimes with data being locked behind proprietary restrictions (Taylor and Schroeder 2015). These data then come to influence how we know disasters and their socio-political foundations, in turn influencing the types of organized responses deemed “appropriate” or “legitimate” (Crawford and Finn 2015).

The “digital divide” thus persists here in numerous ways that delimit the technologies’ potential impact on social change (Graham et al. 2014). Burns (2014) has argued that internal contestations and knowledge politics underwrite the ability to access, leverage, and influence the development of digital humanitarian technologies. For example, installing, customizing, and serving an instance of the Ushahidi\(^2\) crisis mapping platform requires skills in server administration, web scripting, and database management (Brandusescu, Sieber, and Jochems 2015). These limitations parallel Gilbert’s (2010) reconceptualization of the digital divide to account for social capital. In these ways, some are always excluded from humanitarian data (Mulder et al. 2016), but more importantly, asymmetrical power relations are reproduced in the digital humanitarian context (Sandvik et al. 2014; Burns 2015; Duffield 2016). A decade ago, Elwood (2006) argued that framing community organizations’ GIS-based knowledge production as either activist/resistance or as co-optation is problematic, as these roles and relationships are not necessarily singular nor mutually exclusive; these sorts of slippages within digital humanitarianism have not yet been explored.

---

\(^2\) Ushahidi (https://www.ushahidi.com/) is a platform which mobilizes geographically distributed labor to collect, categorize, translate, and georeference data such as social media and Short Message System (SMS) messages.
Still, many remain cautiously optimistic regarding the potential of digital labor for disaster response. Resor (2016) insists on a blurred boundary between “digital humanitarians” and “formal responders,” noting that many digital humanitarians have high degrees of professional experience, and often maintain ties to formal institutions. This upholds decades of research findings about self-organization following crises (Stallings and Quarantelli 1985). Many continue to argue that volunteered geographic information can be useful for generating knowledge of ground conditions following disasters, although its applicability to preparation, mitigation, and recovery are less clear (Heinzelman, Waters, and United States Institute of 2010; Haworth, Whittaker, and Bruce 2016; Shepard et al. 2016).

To date, much research on Big Data tends to understand data as a direct reflection of conditions “in real life,” rather than as reflecting institutional and social contexts (see, for example, Procter, Vis, and Voss 2013). Scholars typically use the term “Big Data” to conjure social media, automatic sensor data, clickstream and web behavior data, and retail purchasing information (Kitchin 2014). Analysis of social media content in particular tends to search for meaning and patterns in data as they are presented (boyd and Crawford 2012). This approach obfuscates the institutional and community-based processes and limitations that frame the types of data produced and the representational strategies espoused (Gray 2012). Geographers have a rich history of interrogating these intersections of data and society (Dalton and Thatcher 2014), and it is in this vein that the current article seeks to address that gap.

This study builds on the existing literature by exploring the ways in which community organizations and individuals engage with digital humanitarian technologies to, on the broadest level, enact social change. By acting on their local and immediate scales, they affect larger scales of data practices and representations. In doing so, they are simultaneously – either consciously or otherwise – acting upon the processes by which marginalization and inequalities are reproduced. Digital technologies here are seen as an interface between multiple actors, including formal disaster response agencies, digital humanitarians, community organizations, and the individual members within these groups. As such, the current project speaks to Barkan and Pulido’s (2017) encouragement to understand the ways claims – for resources, recognition, or participation, for example – are made, disciplined, and addressed. Barkan and Pulido touch on the ways cartographic knowledge production “crystallizes recognition of injustice – even for people not interested” explicitly in injustice (2017, 38). It then becomes important for geographers to question how these claims occur through cartographic visualizations, datasets, social media, and other digital technologies. In what follows I explore these questions in the context of research conducted around Superstorm Sandy, in which digital humanitarians, formal responders, and community organization leaders offered multiple competing interpretations of how needs and knowledge should be captured as data.

Methodology + Context

In late October of 2012, Superstorm Sandy became the largest Atlantic storm on record, its hurricane-force winds stretching over 1,000 miles in diameter (National Weather Service 2013). Most the eastern seaboard of the United States was severely damaged, making this the second-most destructive storm on record to hit the United States (Blake et al. 2013). In New York City, among the hardest-hit areas were the southern Queens neighborhoods on the Rockaway Peninsula (Bloch et al. 2012), although effects were widespread due to flooding, and power and public transportation outages. In the Rockaway Peninsula, however, many ICT networks were disrupted, and economic infrastructures such as the peninsula’s boardwalk and beachside amenities, were destroyed for several years (see Figure 1).
I visited New York City in March 2013 to conduct in-depth semi-structured interviews with local community organizations, digital humanitarians, and senior administrators and policymakers. I focused this research on the Rockaway Peninsula, which was still in initial recovery stages despite the four-month gap and despite containing a wide range of socio-economic statuses (see Figure 2). My visit overlapped with an event hosted by the MoMA PS1 to fundraise for local Sandy recovery, and this event was attended by city councilmembers, key response organizers, and the musician Patti Smith.\(^3\) This research was a component of a larger project primarily located at a public policy research institute in Washington, DC, where for a year I used the extended case method (Burawoy 1998) to understand the societal, policy, and political-economic impacts of digital humanitarianism. In addition to 37 total interviews – of which 7 were in New York, I worked as a participant-observer within several digital humanitarian organizations, and performed archived data retrieval. Interviewees were chosen through a combination of snowball sampling and after identifying key actors as a participant-observer. These semi-structured interviews ranging from half an hour to two hours sought to understand the ways digital humanitarian technologies have shifted data production and representation practices; in New York, the interviews focused in particular on Superstorm Sandy. All these data were collated, transcribed, coded, and analyzed using a discourse analysis framework. In the rest of the article below, I am seeking to understand the social processes that compelled some to enact particular data practices, and likewise I explore the politics and struggles behind those processes; here, communities’ and individual perceptions of the crisis and response efforts are more important than whether or not individual members were factually correct. The data presented below have been selected because they are particularly representative of these processes that appear throughout the corpus of evidence.\(^3\)

---

Figure 1. This photo, taken in March 2013, shows remnants of the destroyed Rockaway Peninsula boardwalk and ongoing beach erosion problems.

Figure 2. The Rockaway Peninsula, where this study took place, has a wide range of socio-economic statuses.

(In)formal Data Collection Practices
Several digital humanitarian groups emerged with disparate data production and representation efforts following Superstorm Sandy, at times to directly confront those of formal responders such as the Federal Emergency Management Agency (FEMA) and New York’s Office of Emergency Management (OEM). Among the various groups at work, multiple competing interests and goals fostered a dynamic characterized by complex actions of resistance, needs provision, and intra-institutional conflicts. At the same time, individuals and groups adopted data production and representation strategies in direct response to the established workflows, relationships between emergency management agencies, and under-represented needs of stakeholders. That is, despite the messy relations between data producers and organizations, the data that emerged to represent disaster knowledge were framed by broad institutional processes.
FEMA, the Red Cross, and city managers in New York began intensive data collection during and immediately following Sandy, guided by institutional normative understanding that the first 72 hours are the most critical for effective disaster response. These efforts sought to produce “situational awareness”: data intended to inform decision-making and resource allocation. Simultaneously, a motley crew of entrepreneurs and technologists gathered under the ad hoc volunteer-based umbrella institution Crisis Commons⁴ to similarly begin producing data and web maps, and begin acting on these knowledge outcomes. According to my interviewee Rachel, a key member of this Crisis Commons group, formal responders distributed the locations of resource centers to potentially affected households on the Rockaway Peninsula. This work marginalized families and households that were bound to their homes and unable to travel to the centers. Rachel and others in the Crisis Commons group believed that there was no formal-sector organized program to reach out to these people, and that without power, these homebound families were “the most vulnerable population.” She and others organized 20 volunteers who visited 500 households in a few hours, to ascertain the needs in the area, which homes had electricity available, and if any medical emergencies were occurring. For Rachel, not only were these data not being produced by formal response agencies, but this was reflective of a limitation more structurally rooted in problematic data practices. Speaking of an organization working with Crisis Commons, Rachel said:

They needed not the same information that FEMA needs. They needed a piece of that information, and they needed it tied to a map in the same way that anyone else did... [But] it evolve[s] into these fiefdoms of power ... FEMA I guess in some places does do door-to-door checking, but they won't share that data with anyone.

Here Rachel points to the ways FEMA’s spatial-informational needs stem from their charter as a formal emergency management agency. Due to issues around privacy and data sharing limitations, on the one hand, and the specific responsibilities which FEMA has been delegated, on the other, the data that the organization produces can be seen to be less useful or accessible to other organizations. Invoking the notion of “fiefdoms of power,” Rachel implies that data sharing can promote empowerment; this line of thinking has been central to digital humanitarianism’s self-marketing, and is here linked to specific data practices.

Rachel and her team took their collected data to formal response agencies to illustrate the value of their digital labor:

They found one woman who was in need of heart and liver (or kidney) medication, and there were several other people who needed to be checked on. We reported that information ... I went to the police station, I talked to FEMA, I talked to Red Cross, and I went to the police station, and finally was talking with the [New York Police Department Chief Information Officer] - I showed him the walk-list, he's like, "Now I get it. I thought you were crazy at first. Now I understand what you're doing." And I was like, “Thank you, this is really important, and I have all the work here and this is the most effective way that you can protect your people. Who can I hand this over to?" And the answer was no one.

⁴ Find more information about Crisis Commons at https://crisiscommons.org/.
What is most important here is the way Rachel and Crisis Commons-affiliated organizations collected their own datasets in direct response to perceived shortcomings of formal responders. Rachel herself does not live in or have significant ties to the Rockaway Peninsula, but instead saw the communities there as in need of her group’s data collection capacity. This created a tripartite relation of power between formal responders, relatively less-affected individuals, and disaster-affected communities. Rachel and her group established new data collection practices without the explicit request of formal responders or communities in the Rockaway Peninsula, instead seeking to address a perceived need. As explained later, such non-residents came to the Rockaway Peninsula en masse intending to help, yet unaware of their own incomplete understanding of the disaster’s socio-economic foundations and impacts. In this way, they were both enacting problematic representations of needs, and organizing to shift resource allocation.

Rachel’s anecdote speaks as well to the politics underwriting knowledge and needs inclusion, and representation as data. Speaking to broader issues of legitimacy, particularly as tied to cartographic representation, in this case digital humanitarians identified gaps in the kinds of knowledges and needs considered legitimate for formal-sector intervention. However, whereas past research has foregrounded the role of maps in these politics, Rachel points to the politics of data itself. Capturing needs as data enrolls a politics of exclusion, as formal actors inadvertently yet by necessity maintain gaps in their collection efforts, and here digital humanitarians contested this politics through new data practices.

Complexities in Institutionalized Spatial Data Production

Oppositional impressions of these digital humanitarian efforts emerged from formal responders and from leaders of other local organizations. These complexities pointed to the multiple competing demands, roles, and needs of organizations, and the data practices they adopt or resist in relation to each of them. Those in charge of disaster-related data collection and production in some ways buttressed the digital humanitarian efforts above, but leaders of other organizations criticized the same efforts. These reactions to digital humanitarian efforts did not themselves foster new data practices, instead providing an additional framework around digital humanitarian efforts for formal responders, and distancing many community organizations from digital humanitarianism writ large.

Harper, a manager for a formal GIS department in New York and advocate of digital humanitarianism, provided spatial data, analysis, and cartographic visualizations for the city during Superstorm Sandy. She expressed concern about the workflows and everyday data practices with which these efforts must align. For formal responders, this consists primarily of the Incident Command System (ICS), an international structured protocol to guide and coordinate emergency response. In an interview she estimated that her department and those with whom she maintains professional connections will draw parallels with extant spatial data platforms when evaluating the potential impact of the above digital humanitarian work. She spoke of the reasons why Rachel and her group likely encountered institutional challenges: “It’s similar to the questions we faced with GIS: where does it fit into the ICS? Unless it gets formal adoption into this framework, it won’t get ‘owned’. Moving a city government takes time.”

Harper asserts that attention to institutional limitations structuring disaster-related data types and data representations comprises an important part of her work as someone whose work directly contributes to the emergency management field. For her, Rachel’s data collection efforts were rejected likely

---

5 See for more details: https://www.fema.gov/national-incident-management-system
because of the unsolicited, amateur nature of the datasets; this contrasts with the purported authoritative, ICS-compliant data production practices of the formal emergency management sector.\textsuperscript{6} The temporal dimension Harper underscores refers to the momentum of current established practices which digital humanitarians confront, and which also frames the kinds of knowledges, needs, and places that are captured as data in disaster contexts.

This sort of reticence toward volunteer data production was echoed by managers in New York’s Office of Emergency Management and the Department of Health and Mental Hygiene. In Superstorm Sandy, both of these agencies produced data for public consumption by publishing information to social media channels but monitored social media only to direct individuals to the proper authorities for addressing their needs. Analytics tools like GeoFeedia helped derive situational awareness from various media channels. These data production and representation techniques stemmed from formal responders’ need to remain within institutional charters and workflow protocols, but in turn generated new knowledge politics as digital humanitarians sought to address what they perceived as the ensuing gaps.

Additionally, within days of the storm, several community organizations were formed on the Rockaway Peninsula to supply labor and resources to underprivileged families in the area. Key organizers for two of these organizations – Rockaway Emergency Plan and Respond & Rebuild\textsuperscript{7} – stated their insistence that their organizations would be directed and organized by people residing in the Rockaway Peninsula in explicit contrast with digital humanitarians such as Crisis Commons.\textsuperscript{8} Rowan, who worked with Respond & Rebuild, expressed to me her frustration that groups such as Crisis Commons had come to the area from distant boroughs in order to represent the communities’ needs. She claimed that these and other digital humanitarians represented needs in ways that elevated Crisis Commons’s mission and marketing over the communities’ need for assistance. To illustrate the politics of needs representation, Rowan recalled a then-recent New York Times article that narrated an outsider’s visit to the stark devastation in Rockaway Peninsula, which quoted the outsider as saying, “I’m driving my big Lexus down here … Thank God the car is dirty” (Nir 2012). Quinn, a leader in the Rockaway Emergency Plan, voiced similar concerns that, regardless of digital humanitarians’ good intentions, they had arrived to an area under duress without a clear commission by the formal responders or by local residents. This established a politics of needs representation that made Quinn uncomfortable, given the vulnerable state of her neighborhood.

According to both Rowan and Quinn, digital humanitarians did not empower locals or accentuate residents’ voices in their efforts. They reached this conclusion through two critiques of approaches toward data. First, they expressed what they claimed was widespread understanding that digital humanitarians did not solicit input from local residents on the types of information that would be collected, or the ways that information would be visualized. Indeed, Respond & Rebuild was established largely to promote local community control over needs representation and needs satisfaction; a primary goal was to gather information excluded by both formal responders and digital humanitarians – information about mold growth after the stormwater surge. Second, organizations broadly under the digital humanitarian umbrella concept arrived to the area to collect and represent data about storm

\textsuperscript{6} Indeed, these sorts of amateur/authoritative binaries have structured most discussions of VGI and the geoweb.

\textsuperscript{7} The official pages for these organizations are: Rockaway Emergency Plan (https://www.facebook.com/rockawayhelp/) and Respond & Rebuild (https://www.respondandrebuild.org/).
damage, but retained control over that data by placing it behind a paywall. Rowan was particularly critical of two organizations of volunteer first-responders, for following this practice. These organizations took this approach largely because they had partnered with a private software company for data collection, and that company had demanded proprietary data retention. Effectively, however, these institutional relationships excluded local communities and organizations from using the data or holding the organizations accountable for their data practices. More broadly, both interviewees expressed their concerns over implications for data ownership and usage.

**Conclusion**

In this article, I have argued that institutional and community-based politics frame the types of data that are produced, and the ways those data are represented, in disaster contexts. These politics stem from the multiple competing data practices enacted by organizations digital and otherwise, formal response agencies, and individuals. Digital humanitarians play an increasingly important role in these politics, as they enroll distributed digital labor and challenge existing practices and workflows. Digital humanitarianism further elicits new data practices, and contestations around how needs and knowledge will be captured as data. These findings illuminate the social and political inequalities of Big Data by foregrounding the struggles and variegations around data production practices. As the research agenda on spatial technologies continues to adapt to technological change, this research suggests geographers should see data not as reflections of on-the-ground conditions, but instead as a representational negotiation rooted in spatial inequalities.

However, spatial technologies hold incredible epistemological and tactical promise, as demonstrated by critical GIS literature (Kwan 2002; Sheppard 2005). To this end, digital humanitarianism is neither limited to nor necessarily characterized by exclusion, and may be leveraged in work toward social justice. Some digital humanitarians use this as a guiding principle in their efforts to fill gaps left by formal response agencies. Yet, the complex politics of needs and knowledge representation I highlighted above insist that digital humanitarians acknowledge and account for these politics, perhaps building technology differently to promote grassroots data production and representation capacity. That is, the history of technology shows digital humanitarians multitudes of ways their technologies can be subverted for productive and positive social change.

To these ends, I conclude by briefly offering two recommendations to digital humanitarians and to policymakers seeking to engage these digital communities and digital technologies. First, as discussed above, data are neither neutral nor direct reflections of on-the-ground conditions, and one should consider not simply data presences and absences but also the contexts and forces which produce discernable spatial patterns. That is to say, content analysis should not take data at face value, but should instead acknowledge the complex processes which lead to some data being produced and not others. Second, and most importantly, in this paper I have identified ways data are always incomplete yet tell an important story. Digital humanitarian technologies hold the potential to facilitate production of data that might have marginal use to formal workflows and institutional structures yet still convey important knowledge about a disaster. For example, digital humanitarian data may capture disrupted interpersonal networks, emotional geographies, spaces of care, new ways of thinking about and relating to urban infrastructure, or communal non-formalized knowledge. It is in these potentialities that digital humanitarianism shows the most promise for social justice work.
Bibliography


