Department of Industrial Engineering and Management Sciences

Northwestern University, Evanston, Illinois, 60208-3119, U.S.A.

Working Paper No. 14-05

Capturing Real-Time Data in Disaster Response Logistics

Kezban Yagci Sokat, Rui Zhou, Irina S. Dolinskaya, Karen Smilowitz

Department of Industrial Engineering and Management Sciences, Northwestern University

Jennifer Chan

Department of EmergencyMedicine, Northwestern University

December 22, 2014

Noname manuscript No.
(will be inserted by the editor)

Capturing Real-Time Data in Disaster Response Logistics

Kezban Yagci Sokat, Rui Zhou, Irina S. Dolinskaya, Karen Smilowitz, Jennifer Chan

Received: date / Accepted: date

Abstract The volume, accuracy, accessibility and level of detail of near real-time data emerging from disaster affected regions continue to significantly improve. Integration of dynamically evolving in-field data is an important, yet often overlooked, component of the humanitarian logistics models. In this paper, we present a case study of near real-time data collection in the days following landfall of Typhoon Haiyan. We detail our first-hand experience of capturing data as the post-disaster response unfolds starting November 10, 2013 to March 31, 2014 and assess the characteristics and evolution of data pertaining to humanitarian logistics modeling. We present a framework for real-time humanitarian logistics data focused on mathematical modeling, which reflects the integration of our recent experience, a survey of different communities producing data and disciplines using data, and a development of measures to evaluate quality of data. Additionally, we provide modeling implications of collected data based on our preliminary data analysis and the proposed framework. The study describes how our humanitarian logistics team approached the emergence of dynamic online data after the disaster and the challenges faced during the collection process, as well as recommendations to address these challenges in the future (when possible) from an academic humanitarian logistics perspective.

Keywords humanitarian logistics · real-time data · classification · logistical modeling · Typhoon Haiyan · disaster

Kezban Yagci Sokat
Department of Industrial Engineering and Management Sciences, Northwestern University
2145 Sheridan Road, Evanston, IL 60208
Tel.: 847-491-3383
Fax: 847-491-8005
E-mail: kezban.yagcisokat@u.northwestern.edu

Rui Zhou
Department of Industrial Engineering and Management Sciences, Northwestern University
2145 Sheridan Road, Evanston, IL 60208

Irina Dolinskaya
Department of Industrial Engineering and Management Sciences, Northwestern University
2145 Sheridan Road, Evanston, IL 60208

Karen Smilowitz
Department of Industrial Engineering and Management Sciences, Northwestern University
2145 Sheridan Road, Evanston, IL 60208

Jennifer Chan
Department of Emergency Medicine, Northwestern University
211 E Ontario Street, Chicago IL 60611

1 Introduction

Over the past two decades, the field of humanitarian logistics has progressed significantly, with a growing number of researchers and practitioners studying problems, such as relief distribution, post disaster debris removal, and evacuation of affected populations. Much work within the academic community has focused on the development and application of operations research tools for humanitarian logistics (e.g., see recent surveys [1,2,3,4]). As the ultimate goals and benefits of these efforts are to improve real-world applications, integration of in-field data is an important, yet often overlooked, component of such humanitarian logistics models. For example, in their recent review, Sangiamkul and Hillegersberg [5] identify only two papers [6,7] out of 30 surveyed that use real-time data in logistical modeling. In another survey, Ortuno et al. [8] describe only two papers [6,7] among 87 that integrate dynamically updated data. Finally, Ozdamar and Ertem [9] acknowledge three papers [6,7,10] out of 110 studies mentioned in their review of humanitarian logistic models, solutions and technologies that capture such data. Outside these academic disciplines, extensive efforts have been made in information communication technology, especially in regards to the use of social media and crowdsourcing in disaster management [11,12]. At the same time, the volume, accuracy, accessibility and level of detail of nearly real-time data emerging from disaster affected regions continue to significantly improve [12]. Considerable efforts are currently focused on collection, aggregation and dissemination of field data, which, together with the help of the humanitarian logistics decision tools, have potential to considerably impact relief efforts [12]. In this paper, we explore the process of retrieving real post-disaster relief data from sources available online and examine the data for the purpose of integrating data streams into response logistics models.

We focus on one specific disaster response, Typhoon Haiyan, as the case study for investigating the role, value and limitations of integrating new information streams for logistical modeling during the aftermath of a disaster. On November 8, 2013, Typhoon Haiyan, named as Typhoon Yolanda, the strongest storm recorded at landfall [13] and one of the strongest tropical cyclones in recorded history [14], hit Philippines and resulted in catastrophic damage throughout the country. As of April 7, 2014, 6,300 individuals were reported dead, 28,689 injured and 1,061 are still missing according to National Disaster Risk Reduction and Management Council (NDRRMC)[15]. Both the Philippine government and international humanitarian organizations started their preparedness activities as early as November 7, 2013 and began response activities immediately following the fall.

In this paper, we present a case study of near real-time data collection in the days following landfall of Typhoon Haiyan. We detail our first-hand experience of capturing data as a post-disaster response unfolds starting November 10, 2013 to March 31, 2014 and assess the characteristics and evolution of data pertaining to humanitarian logistics and our research interests in data for logistical modeling. The case study outlines what data were available to the team, how early data are available, how data change after the disaster, etc. We present a framework for real-time humanitarian logistics data focused on mathematical modeling, which reflects the integration of our recent experience, a survey of different communities producing data and disciplines using data, and a development of measures to evaluate quality of data. We also provide implications for humanitarian logistics models based on our preliminary analysis and the proposed framework.

Typhoon Haiyan represents an evolution of disaster response in which the emergence and growth of data during relief operations bring new opportunities for addressing humanitarian challenges. Multiple factors played a role in the generation of this outstanding level of information including local factors, the nature of the disaster and efforts of the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) and availability of data from Philippine government responding entities. More specifically, Philippines have benefited from advanced information communications technology and widespread media and organizational coverage within the country in the disaster response efforts. Because this case study was performed with an English speaking team, the synergy between information and data predominantly shared in the English language provided an opportunity to pursue this case study. In addition, the Typhoon Haiyan post disaster environment was permissive with respect to information sharing across stakeholders; unlike conflict driven complex humanitarian crises where repressive environments often restrict information sharing, especially through public online sources. Typhoon Haiyan is a sudden onset disaster with relatively predictable timing and location, which made it possible for advance staging of volunteers. OCHA's call for digital volunteer support through the Digital Humanitarian Network prior to the typhoon activated volunteers around the world to participate in collecting and processing information [16]. In addition, the growth of digital humanitarians or "crisis mappers" has expanded non-traditional data streams during recent crises (Haiti earthquake, Pakistan, Chile, Christchurch, Bhopa, super storm Sandy, etc) often in online formats and frequently available to the public.

The rest of the paper is organized as follows. Section 2 provides brief overview of the information and data retrieval process. Section 3 presents the proposed framework for humanitarian logistics data with respect to humanitarian logistics modeling based on the observations from the information retrieval process and findings from the reviews of the information outlets. Section 4 presents the implications for logistics modeling. Section 5 describes the lessons learned and outlines future steps for capturing real-time data in disaster response. Section 6 concludes this paper with final remarks.

2 Case Study: Information and Data Retrieval

This section highlights the chain of events describing our information retrieval process. Figure 1 shows the timeline of this study.

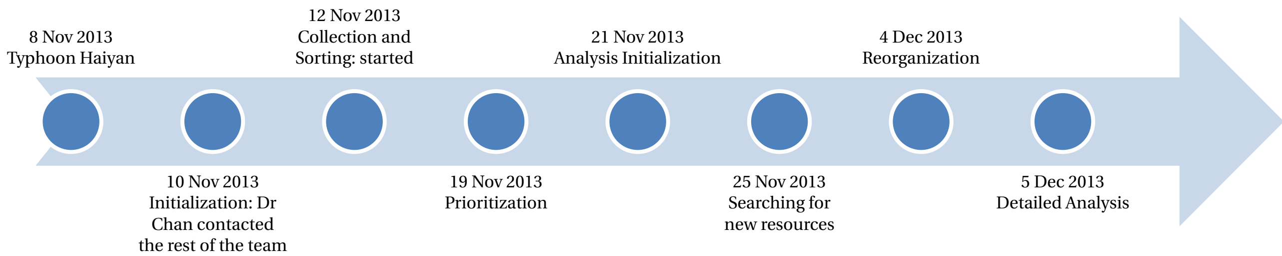

Fig. 1 Information and Data Retrieval Timeline

Initialization. Shortly after November 8, when Typhoon Haiyan moved into the central Philippines region, one of our team members, who is an active researcher and practitioner in the humanitarian technologies community, began receiving numerous email communications from the crisis mappers network [**?**]. These emails contained publicly available website links to information sources pertaining to Typhoon Haiyan. This network is also linked to other digital humanitarian groups that were being activated by the Digital Humanitarian Network (DHN), including the Standby Task Force, Humanitarian OpenStreetMap Team, GISCorps, HumanityRoad, Info4Disaster, MapAction, Translators without Borders, Statistics without Borders and other members of the Digital Humanitarian Network [17].

Collection and sorting. The team recruited an undergraduate research assistant to collect data beginning on November 12, 2013, starting with the resources identified in initial emails. As a first step, we iteratively devised a sorting scheme for data intake. After teasing out the emails relevant to the typhoon, our initial approach in the first days, between November 12, 2013 and November 19, 2013 was to continuously download from the data sources identified on those websites, such as Logistics Cluster and Copernicus Emergency Management Service (CEMS), with the goal of capturing the temporal aspect of the available data. The potential relevance for each data source was assessed using the accompanying description of the data (see Section 3.3 for descriptions). However, since time constraints and our advancing knowledge of humanitarian information sources did not allow for complete understanding of each dataset at the time of retrieval, we continued to download from as many of these sources as we could find, in the hopes that each of those sources might contain desired information.

Prioritization. After one week of sorting data, around November 19, 2013, we realized that the volume, both in number and size, of data was overwhelming. For example, the Humanitarian Response had over a hundred new files [18] and OpenStreetMap had several gigabits of data [19]. The magnitude of the data created a need to balance the tradeoff between processing the known data sources and searching for new sources. After the initial few days of collecting a breadth of data, we attempted to focus on understanding the content of the resources, especially with respect to their logistical content. For almost a week, we focused on differentiating between logistical content, storing and archiving current data according to their content type.

Searching for new sources. As time passed, around November 25, 2013, we continued to search for new data sources. The sources from the email lists served as a starting point, and allowed for retrieval from some important sources earlier than otherwise possible. However, finding new sources required other methods, such as subscriptions to appropriate newsletters and mailing lists, and manual Internet searches. Yet, even with these methods as aids to supplement the manual search for sources, we did not identify all additional sources, particularly those not openly shared on the Internet. This illustrates the challenge to discover the relevant sources fair enough to reduce lost data and the limits of remote research activities that explore field operations. Within this particular case study, many relevant sources, e.g., the OSM repository specific to Typhoon Haiyan, were further researched by our team over one month later, despite OSM/HOT and the American Red Cross activities commencing very early in the response, around November 8, 2013.

Reorganization. With the addition of new resources, the team acquired significant information and continued content differentiation, storing and archiving. On December 4, the team noticed duplication of resources and decided to reorganize the list of resources. At this point, team discontinued downloading from repeating outlets.

Analysis. The team started the analysis of downloaded data on November 21, 2013, which consisted of the initial assessment of the data content, especially the logistical content to be discussed in Section 3.3. While the team worked on analysis and data collection in parallel after this point, the detailed analysis of data sources classifications (see Section 3.3) took upon on December 5, 2013. The details about the analysis efforts are explained in Section 4.

3 Framework

This section presents the development of a framework for real-time humanitarian logistics data for modelers that integrates key aspects of data based on our experience with the Typhoon Haiyan case study. In our framework, we consider three aspects of data from the perspective of humanitarian logistics research 1) who produces data (data gathering *communities*) 2) who uses data (data using *disciplines*) and 3) *measures* that capture quality and applicability of data. These are the three aspects that modelers must understand when assessing new data streams in an effort to improve humanitarian logistical modeling and real-time data use. In Section 3.1, we describe the communities that collect, process and disseminate data through traditional and new data streams, following a disaster. Then, in Section 3.2, we introduce different disciplines that synthesize and integrate available data into their operations according to their specific purpose. Finally, in Section 3.3 we assess the quality and usability of the data and data sources for improved logistical modeling. A more detailed description of the specific data sources used in our study is provided in Appendix.

Figure 2 presents our framework for evaluating new data streams with relationship to humanitarian logistics modeling. We choose to use a circular diagram for our framework to emphasize that the relationship among these aspects is not necessarily linear. For example, disciplines may not only use the data provided by the communities for their own purposes, but also play a role in processing the data. More specifically, researchers studying social media and crowdsourcing may transform information (e.g., identify Facebook or Twitter messages asking for help) into other file formats, which can be useful to other disciplines beyond the initial intended purpose.

3.1 Data Gathering Communities

Multiple communities play a role in the evolution of post disaster data via collection, processing or dissemination. Below, we present some of the key players, based on our observations of publically available online data and information sources between November 10, 2013 and March 31, 2014.

Large International Humanitarian Response Organizations. United Nations Office of Coordination for Humanitarian Affaires (OCHA) aims to plays a critical role in "mobiliz[ing] and coordinat[ing] effective and principled humanitarian action in partnership with national and international actors in order to alleviate human suffering in disasters and emergencies" [20]. In the immediate aftermath of a disaster, when OCHA receives an international call for providing assistance, it often sends United Nations Disaster Assessment and Coordination (UNDAC) teams to provide initial assessment of the

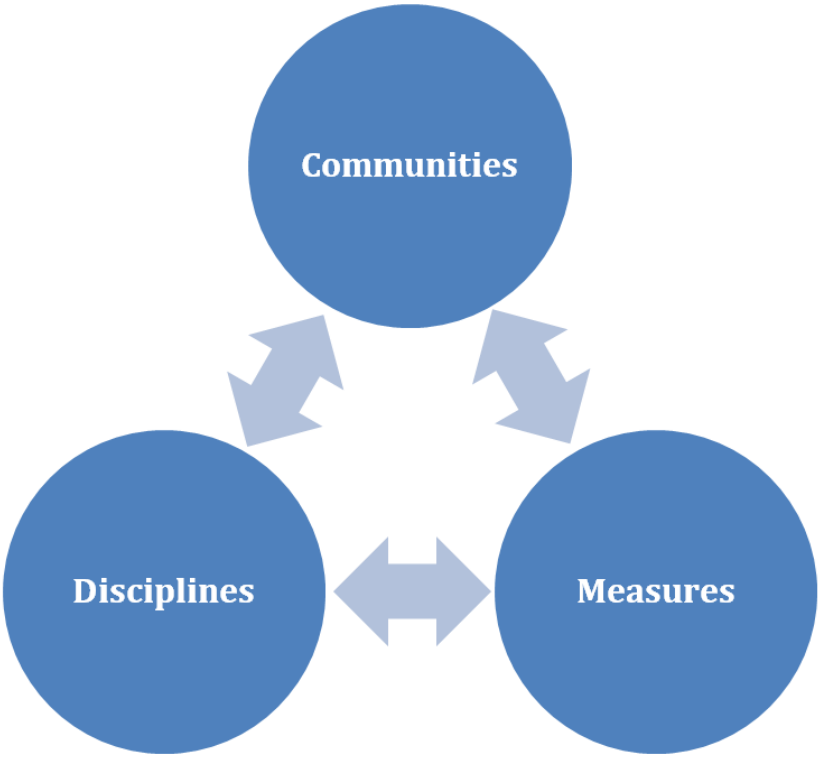

Fig. 2 Framework for Analyzing Real-time Logistics Data for Mathamatical Modeling.

situation [21]. The information management activities within OCHA aim to support the information collection and sharing needs of humanitarian actors to support coordination [22]. In large disasters, responding organizations often coordinate in "clusters" based upon major sectoral activities, such as logistics, health and shelter. In such cases, the information collection, management and sharing are often targeted to the specific cluster's activities. In Typhoon Haiyan, specifically, OCHA facilitated activities of 11 clusters [23].

Information management activities at OCHA take advantage of different types of information in different phases of the response. According to OCHA's description of its services, they aim to create and share information in mediums that are simple to understand and easily accessible. Datasets including common operational datasets, contact lists, and "who, what, where" (3W) data are also maintained and shared by this group. Shared documents in the form of PDF reports and maps are frequently used [22].

Geographic information system (GIS) information plays a key role in OCHA's information management (IM) services. Recent collaborations with external organizations, including MapAction, Humanitarian OpenStreetMap Team (HOT) and GISCorps, have advanced the timely process of map generation. These organizations sometimes leverage microtasking and crowdsourcing methods to process large amounts of geographic information from imagery datasets and non-traditional geographic sources (e.g., satellite imagery, photos) [22,24].

National Government. After each disaster, host government coordinates governmental departments and agencies, such as department of health, department of defense and emergency management authority, for disaster response. In the case of Typhoon Haiyan, The National Disaster Risk Reduction & Management Council (NDRRMC), which functions under Department of National Defense, managed gathering and reporting data [25]. In addition to NDRRMC, Department of Social Welfare and Development (DSWD) played a crucial role in providing information about affected citizens [26]. The detailed description of these two government offices' role in Typhoon Haiyan with respect to data efforts is provided in Appendix.

Digital Humanitarians. In the case of Typhoon Haiyan, OCHA activated the Digital Humanitarian Network immediately after the disaster. According to media reference, this was the first time officials were appointed to coordinate the crowdsourced mapping efforts with volunteer groups [24] during the early stage of the response. Some of these volunteer groups include HOT, Standby Task Force and MapAction. The general role of HOT is to serve as a bridge between the OSM community and the traditional humanitarian relief organizations [27]. In the Philippines, there were more than 1000 OSM volunteers from 82 countries who provided maps to non-governmental organization, including Doctors Without Borders [24] and the American Red Cross [28]. The Standby Task Force analyzed more than one million texts, tweets and other social media posts with the help of MicroMappers software, which uses machine-learning techniques to filter potentially relevant messages [24]. MapAction, a longtime partner of OCHA, worked in the Philippines to generate more than a hundred files per day to be shared with the disaster relief community. With all these efforts, the data from Typhoon Haiyan is a notable example of the evolution of collaboration between digital humanitarian and response agencies, where access to information, collaboration and the next steps of information sharing were pushed forward.

Operationally-focused Humanitarian Practitioners and Responders. Humanitarian practitioners and responders (both local and international) in affected regions are often among the most knowledgeable people about the changing post disaster environment. They are frequently aware of information sources and datasets, sometimes generating data themselves, which may not only reflect the current context, but also represent information and data used by organizations for planning and executing response activities. In our experience, pre-existing relationship of our team member with other practitioners and responders has brought tremendous value to identifing and better understanding various information outlets and how they can be integrated in future logistical models using real-time data.

As information communication technology improves, connecting with responding humanitarian organizations, theoretically, is more feasible. However, developing trusting relations and personal networks still requires years of engagement in working with people from various backgrounds, often different short term goals, but with common overarching missions.

3.2 Data Using Disciplines

Current humanitarian logistics models aim to capture real-time data in order to improve their decision support tools. In this section, we discuss how humanitarian data are used in these models and the assumptions the researchers make, especially in relations to the data availability. Various disciplines utilize humanitarian logistics data and impact their progression. Therefore, the role and purpose of each discipline as it relates to real-time data should be taken into account by the logistics modelers in order to better understand the data characteristics. Thus, we also discuss the role humanitarian data plays for the intended primary users of these data (*practitioners*) and the data collection, processing and communication facilitators (*ICT*).

Academic Humanitarian Logistics. We first describe the current efforts of humanitarian logistics models with real-time data. As mentioned in the Section 1, although the number of models related to humanitarian logistics is growing, models with real-time data are quite limited. We review those papers below.

Liu and He [29] present a decision model for allocation of relief resources in natural disasters using information updates. These updates predominantly contain information on disaster states (population transfer rates) and traffic conditions (road affected level). Authors suggest that these information can be obtained from the disaster database of governmental agencies, such as the National Oceanic and Atmospheric Administration, the National Climatic Data Center, and the National Geophysical Data Center, among others. Liu and He apply their model to Wenchuan Earthquake in China with data provided by China's National Committee on Disaster Reduction.

Sheu [6] develops a dynamic relief-demand management model that forecasts the demand in real-time and dynamically allocates supplies based on those forecasts, as well as urgency and population vulnerability measures. The main components of the information used in the model are 1) time-varying ratio of the estimated number of trapped survivals relative to the local population; 2) population density associated with a given affected region; 3) proportion of the frail population (e.g., children and the elders) relative to the total number of population trapped; 4) time elapsed since the most recent relief arrival; and 5) level of building damage. Sheu uses the official statistics from the 921 earthquake special report from Taiwan to demonstrate application of the developed model. The model contains the most detailed amount of data in comparison to other academic studies we have surveyed and is valuable for estimating regional level demand under dynamic information update. The author also generates simulation data to replace the missing data points in an effort to tackle incomplete information.

Yi and Ozdamar [7] study a dynamic coordination problem of supply distribution and transfer of injured people. They apply their model to an earthquake scenario for Istanbul. The demand distribution (number of wounded people) and supply distribution (people, fleet composition, and total capacity transport) are provided for each time period. Researchers use the widely used data from the Earthquake Engineering Deptartment of Bogazici University [30] for attrition numbers and possible structural damage of Istanbul, which are used to calculate number of affected people. Information about permanent emergency units is gathered from local municipalities and Turkish Medical Doctors Association. However, the information about the number and capacity of vehicles, the capacity of temporary emergency units, as well as how these information are updated were not explicitly provided.

Huang et al. [10] study the impact of incorporating real-time data into disaster relief routing for search and rescue operations in the aftermath of 2010 Haiti earthquake. They use OpenStreetMap to obtain road and building data. They also extract demand information on collapsed structures and trapped persons using Mission 4636, a text message communication initiative. This research provides insights into incorporating crowdsourced data into humanitarian logistics models.

In addition to integration of data into logistical models, some researchers also study classification frameworks for humanitarian data. This work is discussed in more detail in Section 3.3, as it closely relates to our developed measures of data quality and applicability.

Humanitarian Practitioners. Humanitarian practitioners often rely on situational awareness to make critical decisions in difficult situations with limited recourses and time. US Coastal Guard define situational awareness as “the ability to identify, process, and comprehend the critical elements of information about what is happening to the team with regards to the mission” [31]. The availability of timely and accurate data are critical to personnel making operational decisions. There have been numerous and ongoing efforts to improve the collection and management of humanitarian data. The 2002 UNOCHA Symposium on Best Practices in Humanitarian Information Exchange resulted in humanitarian information management principles as accessibility, inclusiveness, inter-operability, accountability, verifiability, relevance, impartiality, humanity, timeliness and sustainability [32]. In the later symposium, reliability, reciprocity, and confidentiality were added to the list [33]. The ongoing Humanitarian Data Exchange (HDX) project, led by OCHA, aims to “make humanitarian data easily available and useful for decision-making,” by bringing together multi-country, multi-sourced, curated data for analytical use through a single platform [34]. As part of HDX project, Humanitarian Exchange Language (HXL) is intended to offer standardization of humanitarian data [35]. In order to further facilitate standardization, the HDE Quality Assurance Framework identifies five dimensions of quality as accuracy, timeliness, accessibility, interpretability and comparability [36].

There are also ongoing efforts among the practitioner to build vocabulary standards for crisis management [37]. This is of particular interest to researchers, as we observe not only differences between researchers and practitioners in the terminology used, but even among various handbooks and guidelines intended for practitioners, we see terminology deviation [38,39,40,41,42,43,44,45]. For example, while humanitarian logistics researchers extensively use the term “supply”, practitioners use “supply”, “resources”, “capacity”, “stockpile” and “availability”. Until these vocabulary standards are developed and implemented in the field, researchers should be aware of the various terms different data sources might use in the same context.

The role that data play in humanitarian operations continues to transform as the data gathering, processing and sharing technologies evolve. Many humanitarian agencies actively acknowledge, assess and forecast the effects of the corresponding changes. For example, the United Nations Foundation Disaster 2.0 report [12] examines the future of information sharing in humanitarian emergencies. The recent annual World Disaster Report 2013 from the International Federation of Red Cross and Red Crescent Societies “examines the profound impact of technological innovations on humanitarian action, how humanitarians employ technology in new and creative ways, and what risks and opportunities may emerge as a result of technological innovations [11].” These and other similar reports can provide logistics models with insights into how data is perceived by the humanitarian practitioners discipline.

Information and Communication Technology (ICT). Information and communication technology plays a critical role in facilitating data collection, processing and communication. From the academic perspective, with the ongoing evolution of these technologies, the studies that analyze their application to crisis and emergency management have also significantly expanded [46,47,48,49,50,51]. Majority of this research focuses on crowdsourcing and social media applications in disaster response [52, 53,54,55,56,57,58,59,60,61,62], with particular interest on Twitter [52]. For example, Ashktorab et al. [61] present a Twitter-mining tool to classify, cluster and extract tweets. The authors include the keywords processed in their study, such as "bridge", "intersection", "evacuation", "impact", "injured" and "damage", among others. The authors implement their algorithm to tweets collected from 12 different crises in the United States. Purohit et al. [60] present machine-learning methods developed for social media specifically to identify needs (demands) and offers (supplies) to facilitate relief coordination, by matching the needs with offers, encompassing shelter, money, clothing, volunteer work, etc. Logistics modelers can benefit and often directly utilize the data collection and processing tools developed by ICT discipline.

3.3 Measures

This section presents the measures that describes the quality of data and data sources, as well as their applicability to different disasters for logistics modeling. Information during the aftermath of a disaster can more frequently be found on websites and is often shared via listservs and emails. Each disaster context will vary by the degree of online information access due to several reasons, such as, type of the disaster (disasters with predictable timing and location, such as hurricanes, versus disasters with unpredictable timing, such as earthquakes), availability of information technology, and level of involvement of host nations and their national and local governments. In order to better assess the quality of numerous information sources that emerge after a given disaster and applicability to other disasters, we classify the outlets (see their descriptions in Appendix) and data provided from these outlets based on a number of measures relevant to the focus of this study. As humanitarian logistics modelers, we first determine broad areas of quality and applicability measures, which help us understand humanitarian logistics focused real-time data and their indication for modeling. We then appoint attributes of data and data sources to explain each measure in detail. (See Figure 3.)

The importance of overall quality measures and the level of their applicability depend on the purpose of the data user (e.g., research, situational awareness and decision making on the ground); thus we span these measures around user's *purpose*. The data format, which is a critical attribute for our study, is highly correlated with the purpose of the research team. Data format refers to the format of the files that data or the information relevant to our humanitarian logistics modeling focus is represented. The data is available in many formats from portable data formats (PDF) to keyhole markup language (KML) and Microsoft Word Documents (.doc). The format of the information disseminated through various sources is an important characteristic, especially for its integration into logistical models. See Section 4 for more discussion.

We survey the data standards and different disciplines described in Section 3.2 and propose four measures: 1) *relevance*, 2) *timeliness*, 3) *generalizability*, and 4) *accuracy*. Figure 3 summarizes the applicability measures studied in this paper. Below we describe the measures and attributes used in our framework based on their utility for logistical modeling and potential challenges. Details about the modeling implications are discussed later in Section 4.

The framework of data and data sources listed below is an amalgamation of different data attributes for humanitarian logistics modelers. We should note that while some attributes listed below might refer to multiple measures (e.g., the classification category primary purpose might infer about the relevance and generalazibilty), some attributes can be related (e.g., local factors and Logistical Performance Index). Table 1 in Appendix provides an example of classification of the information outlets and data available at these resources following Typhoon Haiyan using the attributes below.

3.3.1 Relevance

Relevance is determined by whether the data meets the needs of its users [36]. Relevance of the data obtained from the data streams refer to the degree to which the data meet the current and future needs

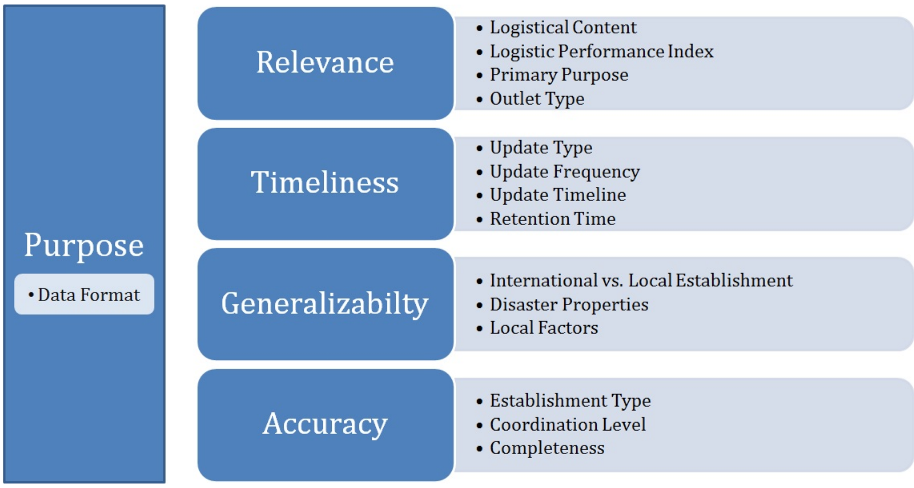

Fig. 3 Applicability Measures and Related Attributes.

of the data required for logistical modeling. For our framework, we identify the following measures that represent relevance.

Logistical Content

Humanitarian and relief organizations constantly collect, process and disseminate immense amounts of information in a broad range of settings and applications. Our work and this specific paper focuses on humanitarian logistic operations in the post-disaster settings, such as on-the-ground operations immediately following natural or man-made disasters (e.g., immediate medical assistance or search and rescue operations). Our first step consists of identifying specific types of information relevant to humanitarian logistics. The data are divided into following categories: demand, supply and infrastructure. These categories are similar to those used in literature [3,63].

– Demand: Effective and efficient relief efforts require the identification of the location, quantity and types of supplies needed within the affected region. Demand in these settings can correspond to needed physical goods, such as food, medication or shelter, as well as needed services, such as medical assistance, rescue, and telecommunication.
– Supply: Information about relief supplies that are pre-existing or gathered after a disaster, transportation vehicles and expert or key personnel (e.g., see [64] for examples specific to Typhoon Haiyan) is another important component for the relief efforts.
– Infrastructure: In order to facilitate distribution of supplies to demand, we need to have knowledge of the infrastructure (e.g., roads, airports, seaports and their post-disaster conditions). While this critical component can be highly uncertain, there is great potential benefit from accurate and timely data from the field.

Logistical Performance Index

The World Bank Logistics Performance Index (LPI) measure the "friendliness" of a country based on six factors: customs, infrastructure, services quality, timeliness, international shipments and tracking/tracing [65]. Tatham et al. [66] suggests Logistics Performance Index as one of the four factors impacting the logistical preparation and response. LPI indicates a disaster's logistical relevance for modeling. Moreover, LPI is also related to the local factors under generalizability (of data and data sources), which will be discussed later.

Primary Purpose

Primary purpose indicates the focus of the information posted in the outlet and/or role of the organization providing the information. Some of the examples are initial assessment, evacuation and providing maps. The alignment of data source primary purpose with the goals of logistics modelers impacts the data relevance.

Outlet Type

Outlet type identifies the ownership of information, as observed in our study. For this specific disaster, we distinguish between two types of information outlets: original information outlets and aggregator information outlets. Original information outlets refer to organizational websites that primarily provide information and data either collected by that organization or transform the data for their specific response purposes. We use the term aggregator for outlets that predominantly disseminate information collected by various other sources.

3.3.2 Timeliness

The timeliness of data refers to time dimensions of the collected, analyzed and disseminated data. Timeliness is deemed by many users as generally the most important characteristic of data [67]. Humanitarian data should be collected, analyzed and disseminated efficiently, and must be kept up to date [32]. Timeliness is also defined as the delay between when the data are collected and when data become available and accessible [36].

Data Update Type (new update/incremental vs. overwrite)
Data update type represents the method by which the status updates are provided after the initial file upload. Incremental updates suggest that the new information is described in a new file. On the other hand, overwrite updates indicate that the additional information is appended to the existing file containing the original information.

Update Frequency
Update frequency represents the frequency with which data are updated. Update frequency can be by minute, hour, day or other. Due to the nature of humanitarian operations and impact of time in the output, the humanitarian community benefits from data being updated timely and frequently.

Data Update Timeline and Retention Time
While both data update timeline and retention time are associated with the time perspective of data, update timeline refers to the timeline from the initial time the data are started to be uploaded until the last time data are uploaded. On the other hand, retention time captures the last time when the data will be available for public use. In other words, update timeline describes when files are updated on the website, e.g., from November 1, 2014 to December 15, 2014. Retention time is for how long that data will stay up on the website, e.g., the data can be accessed for another year after it has been uploaded. Retention time is also often associated with the establishment type of the information outlet, which is discussed later.

3.3.3 Generalizibility

Generalizability measure in this framework indicates the applicability of the information obtained from the data resources of a particular disaster to other disasters for preparedness, analysis, lessons learned and evaluation. We determine establishment's local vs. global designation, disaster properties and local factors as key indicators of generalizability.

Local/ National vs. Global/International
This classification addresses whether the information source is administered by an international organization or a local government/organization where the disaster occurs. It signifies the level of involvement from local government in the disaster response operations.

Disaster Properties
As the name suggests, this attribute describes the main characteristics of a disaster. Tatham et al. [66] develop a 13-parameter framework that captures the factors impacting logistical preparedness and response. Significant part of the classification categories from Tatham et al.'s framework, such as the time avaliable for action (disaster onset), disaster size, magnitude of impact, duration of time and environmental factors (such as the topography or weather conditions of the affected area) are related to disaster characteristics used in our framework.

Local Factors
Similar to LPI, this measure indicates the local characteristics of the area where disaster occurred. While LPI focuses on factors that impact logistical performance directly such as infrastructure, local factors refer to metrics for local environment. L'Hermitte et al. [68] present a classification model of disasters from humanitarian logistics perspective. Their model composes the time and space components of the disasters and five external situational factors of the disaster environment. The external factors stated in the paper are the government situational factors, the socio-economic situational factors, the infrastructure situational factors, the environmental situational factors, and the conflict environment. Five external situational factors of the disaster environment of L'Hermitte et al.'s work [68] and some of the categories

from Tatham et al's 13-parameter framework [68], such as geographic context, population density, per capital GDP, and potential for reoccurence of the disaster in the same area are examples of local factors. Another local factor is the language of the local environment as in the case of 2010 Haiti earthquake [12].

3.3.4 Accuracy

Multiple researchers discuss data accuracy directly or using related terms, such as reliability, verifiability and accountability [32,33,34,36,69]. According to the Humanitarian Data Exchange Quality Assurance Framework, the accuracy of the data is defined as “"the degree to which the information correctly describes the phenomenon it was designed to measure” [36]. Synthesizing the definitions from these resources, we define accuracy as measure that represents the reliability and the credibility of the information obtained from the data streams and data sources.

Establishment Type
Establishment type denotes whether the placement of the data (e.g., website or repository) was established specifically and solely for the purpose of a specific disaster or maintains data across multiple disasters. We use establishment type to separate data sources into two categories: multiple disaster and disaster specific. *Multiple disasters sources*, e.g., ReliefWeb, correspond to organizations and websites involved with data on disasters prior to the given disaster, often retaining data from such disasters. Such sources or repositories usually retain data for multiple disasters after the relief operations are completed. *Disaster-specific sources* generally provide information during relief operations of a specific disaster and may become unavailable shortly afterwards.

Coordination Level
This attribute represents the level of coordination among different communities during disaster response. It might indicate various types of collaboration, such as coordination between different relief agencies and coordination between local and international governing bodies. The concept of coordination through subgroups introduced by Jahre and Jensen [70] aims to organize humanitarian help with a number of different areas by predefined management. Jahre and Jensen discuss the importance of the balance between horizontal and vertical coordination in any period of disaster management. Additionally, the authors mention the role logistics cluster, one of the 11 OCHA formed clusters, on information management and challenges of coordinating the information.

Completeness
Compeleteness represent whether there is missing information or not. Examples of incomplete information might be in regards to status of certain roads or damage level of buildings.

4 Data Analysis and Implications for Disaster Response Logistics Models

Using the applicability and quality measures that we present as part of our framework for real-time humanitarian logistics data (Section 3.3), we identify the following implications of the analyzed data for logistics models. We also provide examples of data availability analysis while discussing implications. As we continue to learn more about humanitarian organizations, datasets and information sources and apply them to the process of integrating near real-time data into humanitarian logistics models, we anticipate our initial findings described below may evolve.

Data Format. Data format plays a critical role in *accessibility of data* and *smooth integration of data into models*. Many of the available files identified in the case study depict humanitarian logistics information such as roads and hospitals, but frequently in static formats. Optimal data formats to import into logistics models would be editable documents or dataset file types. For example, PDF maps contain great information about severity of damage in an area; however, due to their format, it is hard to access information about road structure. Many released written reports do not contain easily or quickly transferable numerical data. KML files and SHP (shape) files might contain the relevant data for testing models, but not in an immediately accessible manner, often requiring file conversions and format manipulations. On the other hand, even after these files are converted, discrepancies may exist between the content. Two different sources converted through the same online SHP to CSV formatter [76] can return different marking systems, e.g., XY coordinates vs. “osm_id,” as location indicators. This variation may be problematic for analysis, since it might require multiple different infrastructure information sources. On the other hand, some sources contain Microsoft Excel data spreadsheets, and classification in these data simplifies the identification process of sources with promising data for the purpose of optimization models.

Logistical Content. As discussed in Section 3.2, there are differences in the terminology between researchers and practitioners. In order to find the *relevant* information (e.g., *demand* and *infrastructure*) to integrate into models, academic researchers should first learn about these differences. For example, in the case of Typhoon Haiyan, while searching for relevant information, we establish a set of key search terms that are used to identify the potential data sources within our compiled database that would contain information about service *demand* and *infrastructure*. We use sources from different disciplines described in Section 3.2, to help us build the list of these search terms. Then, using this list, we examine the data gathered from sources such as OpenStreetMap, Logistics Cluster, NDRRMC, and Copernicus EMS for availability of the logistical content (see Appendix for description of these data sources). The filtered data are then analyzed to identify the specific content they contain in relation to the kind of information one would need for logistical models. For example, we want to know the level of damage each road link sustained, its post-disaster status and location of potential beneficiaries of medical and rescue services.

Logistical Performance Index. This measure assigns a numerical value capturing the performance of a country's logistical structure. According to existing studies (e.g., [77]), as the value of LPI increases, the expected number of affected people generally decreases. Logistics modelers can benefit from this information when estimating service demand for a given region. LPI can also be associated with *generalizibility* of the models, since using one country's LPI value, modelers can evaluate applicability of the available data from one country to another. Furthermore, higher LPI scores usually correspond to less restriction from the government [78] on relief response operations, corresponding to another *generalizibility* measure of the data.

Primary Purpose. Primary purpose of the data outlet or organization might help researchers search for *relevant* information for their models. For example, if a researcher is working on search-and-rescue operations, it might be easier for him or her to start from an outlet that is focused on initial damage assessments. More specifically, in the case of Typhoon Haiyan, a researcher can begin his or her analysis from Copernicus EMS, NDRRMC and OpenStreetMap (see Appendix for primary purpose of surveyed information outlets in this study).

Outlet Type. This type of classification can be helpful for practitioners and researchers for the following reasons. Original information outlets might be a good choice to look for when a researcher or practitioner is looking for a certain type and/or format of information, such as maps or reports about damaged areas. In this case, researchers or practitioners might need to search for several original information outlets to find particular data or accumulate a series of different information pieces. Aggregator information outlets generally compile information about supply, demand and/or infrastructure from assorted outlets. Thus, one might want to start their search from these sources. However, researchers should be aware of duplications due to multiple posting of either same documents or segments of data, such as the NDRRMC situation reports placed on both, ReliefWeb.int and the NDRRMC site.

Update Type. Practitioners and researchers may have different needs with respect to data updates. While the practitioners may seek the most current data with cumulative statistics for operational decision-making during a certain disaster, logistics modelers prefer to see the evolution of the post-disaster data. As a result, while an overwrite update may be preferable by practitioners, it is not desirable for humanitarian logistics researchers who focus on adaptive modeling. Modelers generally find the piecewise information about additional available roads or estimated needs (e.g., refugee camp populations requiring food from WFP) at a location more useful. Update type also provides suggestion about prioritization of data collection process. Sources such as OpenStreetMap may have openly available data that archives the prospective near real-time updates of roads, while other maps that have overwrite updates should be monitored frequently to capture the evolution of data.

Update Frequency. Update frequency may have large influence on modeling decisions. In this case study, many organizations that share data sources appear to update and upload their datasets daily and frequently post new content data related to logistics. This may influence the model type and the inclusion of the dynamic information into the models. For example, as the frequency of the information increases, a researcher might prefer dynamic programming to multi-stage optimization for modeling.

Update Timeline. As stated before, *timeliness* is generally one of the most important measures of data, and it has several implications for modelers. Different timelines might express different value to various types of modeling purposes. The first available post-disaster data is crucial, especially when a researcher wants to find as much information as possible to understand the immediate context and link information with high priority humanitarian logistics activities within a certain period after the disaster. For example, the data available during *golden time* (first few days after the disaster) are vital for modelers focused on search-and-rescue operations [10]. While searching for the logistical content, we restrict our initial data analysis to immediate post-disaster response operations, such as search-and-rescue and immediate medical assistance, therefore we present the study that targets the data from the first few weeks following the typhoon. As the initial few days pass after the disaster, the information about the supplies (which team is where with how much medical or other supply) becomes widely available. This information provides a basis for the relief distribution modeling. Additionally, longer timelines imply more data for the modelers, which allows them to conduct more comprehensive analysis for test case generations.

Retention Time. Similar to update type, retention time also might impact and aid in enhancing the data collection process. Longer retention times enables researchers to access time sensitive critical information. Postponing the collection of data that remain well after their posting date may allow for greater time spent on more volatile sources.

Local/National vs. International. While establishment type and primary purpose might inform practitioners about the reliability of the data and other extends, local/national vs. international data ownership might inform about the *generalizability* of this data and analysis for future disasters. For example, the level of involvement from the local government for post disaster response can be included in the discussion of different disaster comparisons. Similarly, depending on the disaster type, pre-disaster evacuation efforts of the local/national government can be a critical factor when comparing different disasters and making inference from them.

Disaster Characteristics. Disaster characteristics can inform modelers about decision making in different stages of the disaster cycle. For example, a sudden onset disaster with predictable time, such as Typhoon Haiyan, can help modelers and practitioners on the ground to improve the prepositioning strategy to save as many lives as possible. In addition, as in the case of Typhoon Haiyan, disasters with predictable timing impact the level of information available in the response phase by advance notice, which can impact the specific characteristics of the model and model validation.

Local Factors. Geographical context of the local area where disaster takes place might provide multiple insights for modeling. For example, Philippines being a combination of islands implies routing decisions using different modes of transportation, as well as coordination of relief items among islands. Furthermore, the fact whether or not a given island is the hub for relief operations can also impact the routing decisions. Another local factor, disaster reoccurence probability, can provide useful information to logistics modelers at various stages of the disaster management cycle. For example, Ergun et al. [79] describe numerous efforts of Waffle House Restaurants to effectively respond to hurricanes. These efforts include equipment prepositioning, special menus and advanced personnel scheduling. Similar strategies can improve disaster management in areas that are prone to storms such as Philippines.

Establishment Type. Establishment type can often inform researchers and practitioners about *reliability, completeness* and *accuracy of information.* Establishment type often relates to the structure of the organization that maintains the information outlet. Multiple disaster data sources may possess additional verification processes, which may increase their reliability. However, these data sources (e.g., PDF maps) rarely contain the raw datasets behind them. In contrast, based upon this specific case study experience, we observed that disaster-specific sources may directly post datasets without progressing through a verification process, or may not fully share the verification process with the public. They may also lack the level of reliability and trust in comparison to sources established for previous disasters. Regardless of the situation, connecting with practitioners at appropriate times might assist in further understanding of how and whether data are verified. It may also open up opportunities to explore datasets in pre-filtered formats, reliable short-term sources, or data sources more relevant to practitioners. However, data sets may not necessarily be complete or entirely accurate, requiring data fusion or synthesis of various datasets together to achieve logistics modeling requirements. The resultant fused datasets will need to be reassessed for applicability for effective on-ground operations. Overall, modelers might need to account for the reliability, completeness and accuracy of data in their applications.

Coordination Level. Examples of coordination after a disaster can be coordination within a cluster, among the clusters, and between local government and international efforts. Key to a successful coordination is exchange of information, which in tern results in additional information generation (e.g., cluster reports to be shared with participants). In addition, the involvement of various parties in the combined mission improves data *accuracy* as information is validated by the distinct participants. Moreover, similar to the establishment's international vs. local designation, the level of cooperation of the local government with the international communities in the disaster response efforts impacts the evolution of available data.

Completeness. Missing data raises questions regarding *accuracy* of information to be used in the logistical models. In order to account for the inaccurate information, uncertainty factors should be included into the models. For example, when searching for logistical content that is of benefit for search-and-rescue operations, we first conduct data availability analysis for road damage. We look specifically at the data for Tacloban City, which was anticipated to have more data in comparison to rural areas due to various factors, including the operational focus of response toward this region as well as the urban context [71]. The number of roads labeled with some level of damage and the total number of roads recorded were compiled for each day between November 7, 2014 and November 28, 2014. However, even by November 28, 2014, only approximately 5% of the added road links among studied data sources (OpenStreetMap, NDRRMC, Copernicus EMS and Logistical Cluster) were labeled to contain some level of damage. This finding raises the question of whether the available information is adequate to understand the full picture and build appropriate models that resemble real-life setting.

In order to obtain service demand information for logistical models, we next examine the available building data, which might inform us about the individuals either trapped under collapsed structures or displaced persons due to loss of property. For the building data, we focus on five cities: Tacloban City, Guiuan, Palo, Ormoc and Cebu. Notably within three of these cities (Tacloban City, Guiuan and Palo), the percentage of buildings with "collapse" or "damage" indicators ranged between 40% and 60% by November 20, 2014. Even though the damage percentages were high in those three cities, the damage was not reported at all in two of the cities, Cebu City and Ormoc. Moreover, the data did not even appear in Palo and Tacloban City until the 14th, and in Guiuan until the 19th. Finding that up to 60% of buildings are damaged raises doubts about the *accurate* data collection and its *relevance* to decision support tools and requires additional analysis.

5 Lessons Learned

This study enables us to better understand the current situation of real-time data, how data evolve, and to what extent real-time data are available. We believe that this valuable experience will inform and aid modelers in building improved models. Section 5.1 describes the challenges faced during the study, and recommendations about how these challenges can be addressed in the future (when possible) from academic humanitarian logistics perspective. Section 5.2 addresses the role that well established relationship between practitioners and researchers play in data evolution and the applicability to logistical models.

5.1 Challenges Faced During the Collection Process and Potential Solutions

Time sensitive information. As information evolves after a disaster, some sources that commonly recur during separate disaster relief efforts do not retain their data for long periods of time. This generally depends on the *establishment type* and *update type* of the information outlet. Investigating *retention time* and *update type* of information outlets before the onset of a disaster can aid in gathering time sensitive information. For instance, collection and analysis of data from resources that have the tendency to retain their files longer can be postponed to later times depending on the ultimate goals of the data. If a researcher focuses on modeling the initial few days of the disaster, this approach might not work. However, if a researcher wants to model a later period, postponing collection of retained files can be beneficial, thus avoiding the trade-off between collection of time sensitive information and prioritization of analysis. Additionally, *format* of the time sensitive information provided by an outlet tends to be similar for each disaster. Familiarity with the *data format* can ease the data collection process. Moreover, some outlets, such as OSM, benefit from specifically searching for a separate repository that they might link to on their wiki webpage. Accessing those repositories in the early days of disaster response supports smooth collection of time sensitive data.

Delays. The data on collapsed and damaged buildings did not begin to appear in Palo and Tacloban City until November 14, 2013, and in Guiuan until November 19, 2013. Depending on the main objectives of the humanitarian agencies using these data, those dates might be too late. For example, from the perspective of search-and-rescue operations, receiving information three days after a disaster may seem too late to assist most of the people trapped under buildings. However, people who were stranded but not directly impacted by the collapsed structures would likely still be alive during that particular time frame and could benefit from relief supplies.

Data explosion and information overload. With the emergence of technology and increasing number of humanitarian organizations, the amount of information available after a disaster is accelerating [12]. From the perspective of a humanitarian logistics research team, it is challenging to account for this information load in a timely manner. One major factor is a limited research workforce. Limited human resource capacity for information processing is a shared challenge across the humanitarian ecosystem and acutely experienced by field practitioners. The evolution and success of many digital humanitarian efforts is the harnessing of remote workforces, often through crowdsourcing and microtasking efforts. Based on this case study and our team's experience, a possible solution for this issue is to recruit additional help in the data collection process whenever possible. Exploring collaborative opportunities with research teams and potentially practitioners to build a feasible and appropriate workforce to identify, filter, assign and prioritize humanitarian logistics datasets for modeling purposes may be a long-term goal. In the short term, future efforts might consist of two groups working in tandem on the data retrieval: one can search for new sources and initialize their retrieval, while the other investigates whether to continue retrieving data from the identified sources or not.

Duplication. Further complications arose when we observed that several data sources were reposting data from other sources, and many updating files with recent time-stamps appeared to be identical to older versions of the files. Recognizing the overlapping data segments between *primary information outlets* and *aggregator information outlets* early is a key point for researchers. Furthermore, prior information about the expected *update timeline* and *primary purpose* of the information outlets can help resolve these issues. For example, if an information outlet is only focused on initial damage assessment, a researcher might stop downloading from this outlet a few weeks after a disaster to prevent any possible duplication.

Relevance. Some of the data retrieved were not as relevant to humanitarian logistics models as originally hoped. Identifying *logistical content* of an information outlet and prioritizing these outlets might be helpful in organizing data collection process in the future. For example, seeking out resources, such as HOT/OSM and the Logistics Cluster, earlier in the future disasters can be valuable since the coordinates used in OSM provide information about the damaged structures. Additionally, collaborating and supporting these organizations before disasters strike can help researchers to understand data relevance more clearly.

Compatibility. Even assuming file compatibility, problems might exist between perceptions of the sources and how the sources are developed. For example, when one data source is developed by people on the ground, and another source is developed by digital humanitarian mappers, conflicting information would have to have a system for prioritization. Such system would also require differentiation of information obtained from mappers versus in-filed personnel. Furthermore, the particular mapping techniques of various sources may differ. One information source may mark a singular node as damaged or impassable, whereas another source might tag the entire street. This discrepancy can cause major differences in routing decisions and make it challenging to combine multiple resources to build a larger database.

Availability. Our data analysis shows that there is only information for 5% of the roads that indicates some sort of damage for road structure. This low level of information in the case of this high impact disaster, with high level of media coverage and large amount of data tracking efforts within the first few weeks of the typhoon landed show that the available information is not enough to integrate real-time data into models without putting efforts into adding accuracy measures and finding missing data. On the other hand, 60% of the available building damage information also requires validity check for source data.

Availability of data across regions change. Some locations receive more attention than others. In particular with the building data, Palo, Tacloban City, and Guiuan appeared to receive more mapping than Ormoc or Cebu City. While some of this can be attributed to proximity to the storm at the peak intensity, looking at displacements suggests that more people were affected in Cebu and Iloilo than Guiuan, yet these regions were less mapped [74]. Some of these variations in coverage may be due to the focus and collaboration of mapping activities with the online communities such as OSM/HOT, and need to be further explored. We also observe that most of the damage indicated by the data was for roads along the coast. A possible explanation of this phenomena can be the presence of multiple medical teams close to the coast [72], as well as OSOCC (UN On-Site Operations Coordination Centers) predominant locations along the coast [73]. In addition, digital humanitarians assisting with updating maps might more easily distinguish a damage for a larger coastal road than more crowded neighborhood streets.

We recognize that numerous pieces of data from aggregator outlets have citations to their primary information sources. However, availability of the underlying raw data is frequently limited for a number of reasons. In addition, multiple primary information outlets share PDF maps, yet the detailed information about the infrastructure damage is challenging to obtain since the original core datasets from the primary source are infrequently cited or made available. This results in information loss.

Technological Status of Disaster Affected Regions. The pre- and post-disaster states of the communication system play a significant role in the opportunities, limitations and gaps of the available data. No matter how technologically advanced a particular geographic area may be, the gaps in telecommunication coverage in the post-disaster setting are often present. Significant communication problems had risen due to the destruction of power and communication lines in Philippines soon after the Typhoon Haiyan [71]. Over a month after the onset, connecting with field teams within specific regions on a daily basis presented significant challenges, as exemplified by "a survey undertaken in the affected community in Guiuan which reconfirmed the need for clearer and more frequent communication between aid partners and affected communities" [91]. Recognizing the damage sustained by regional communication systems can help researchers understand the information flow and better explain missing data (*completeness*) for specific geographic regions and time periods. Absence of information flow from an area can also serve as a signal of significant damage and imply increased needs for humanitarian relief (i.e., *demand*).

5.2 Prior Relationships and Collaboration between Practitioners and Humanitarian Logistics Researchers

Prior Relationships. The considerable amount of disaster relief operations relevant data came to one of our team members' attention, a humanitarian practitioner and responder, shortly after the Typhoon Haiyan through her pre-existing relationships with other humanitarian practitioners and responders. These prior connections aided our team in identifying and exploring various information sources and potentially better understanding the value of the data and information for decision making purposes [12].

Retrieval of information needed for humanitarian logistics models and knowledge of in-field data collection and dissemination come from a unique collaboration between logistical researchers and humanitarians. Establishing such collaborations prior to the onset of a disaster increases the success of such endeavors. Well formed and growing relationships allow for parties to gain insight into each other's respective terminologies and broader domains. Such insight may enable each respective party to alert the other about potential opportunities for exploration, such as the uniqueness of Typhoon Haiyan with regards to public data. Developing relationships to such a point may require years of collaborative work on similar planes, and an openness to take risks in "thinking outside the box" of our respective fields and disciplines. The data retrieval and aggregation process described in this paper would not have been possible to carry out in a timely fashion without the pre-existing relationships between researchers and humanitarian practitioners.

Role of Humanitarian Practitioners. Networking with practitioners plays a significant role in learning about real-time data and implications for humanitarian logistics research. While raw datasets may be optimal for logistics modeling, access to these datasets may require a trusted understanding among practitioners, their respective humanitarian organizations and a research group. The agency's ability to understand the value of the research for their organization or sector will likely be one of many requirements for data sharing of non-public sources; rather than just the academic value of the research itself. Researchers with humanitarian experience or teams who invest in ongoing relationship with organizations will hopefully build common bridges for future collaborations in research, training and practice.

Similar to developing research collaborations, trust networks between humanitarian relief experts within the researcher groups and in-field practitioners should exist prior to the onset of a disaster or between disasters. These times will allow both groups to learn from one another and share insights when immediate response or early recovery activities are not ongoing.

Role of Humanitarian Logistics Researchers. Humanitarian logistics researchers can contribute to helping the humanitarian logistics community identify the key information needed for successful logistical operations. Through comprehensive mathematical models built specifically for the emerging data sources, researchers can identify the most valuable and promising data for the purpose of more efficient humanitarian logistics operations, and ways to integrate these data into a decision making process. Ideally, validated humanitarian logistic models generated from near real-time data shared from humanitarian agencies would undergo a series of iterative processes with practitioners to translate logistic models into accessible field-appropriate tools for field logisticians or agencies to assist in their operational activities.

6 Conclusion

This study presents a first-hand experience of collection, processing and analysis of near real-time information for humanitarian logistics models in the case of Typhoon Haiyan. Keeping the case study experience in mind, we survey different communities producing data and the disciplines processing and integrating data to their operations. We define a set of measures to assess the quality of the data and their applicability to different disasters. We merge these efforts under a framework for real-time data pertaining to humanitarian logistics models. Additionally, we provide modeling implications of data based on preliminary data analysis. The study describes how our humanitarian logistics team approached the emergence of data after the disaster and the challenges faced during the collection process, as well as our observations.

The information and observations included in this study are based only on one disaster, Typhoon Haiyan. Future experiences might differ based on multiple factors, such as the disaster type (e.g., complex emergency, man-made disaster), ICT environment, and involvement of organizations and affected populations. The information outlets described and analyzed in this work constitute only a subset of the available resources and focus on those with English content and online availability. The description of organizations and digital humanitarian groups involved in information management and data sharing is based upon a growing knowledge of our research team and one that is a work in progress. Furthermore, the classification provided in this paper is only one of the many possible ways, where other researchers might approach the same data differently.

To the best of our knowledge, this is the first study conducted by humanitarian logistics researchers focusing on real-time data collection process in post disaster setting. It also presents a unique team approach that combines the expertise of both humanitarian logistics researchers and a researcher with humanitarian practitioner experience. The study enlightens researchers about the availability of real-time data and its challenges. Additionally, it provides a ground work for integration of real-time data into logistical models.

Acknowledgements This work has been in part funded by the National Science Foundation, Grant CMMI-1265786: "Advancing Dynamic Relief Response: Integration of New Data Streams and Routing Models" and accompanying REU supplement, CMMI-1265786/001.

References

1. Altay, N., Green, W.G. III (2006). OR/MS Research in Disaster Operations Management. *European Journal of Operational Research*, 175(1), 475-493.
2. Kovacs, G., Spens, K.M. (2007) Humanitarian Logistics in Disaster Relief Operations. *International Journal of Physical Distribution* & *Logistics Management*, 37(2), 99-114.
3. Ergun, O., Karakus, G., Keskinocak, P., Swann, J. Villarreal, Monica (2010). Operations Research to Improve Disaster Supply Chain Management. *Wiley Encyclopedia of Operations Research and Management Science*, doi: 10.1002/9780470400531.eorms0604
4. De la Torre, L.E., Dolinskaya, I.S. and Smilowitz, K. (2012). Disaster Relief Routing: Integrating Research and Practice. *Socio-Economic Planning Sciences*, 46(1), 88-97.
5. Sangiamkul, Esther and Hillegersberg, Jos van (2011) Research Directions in Information Systems for Humanitarian Logistics. In: 8th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2011, 8-11 May 2011, Lisbon, Portugal.
6. Sheu, J.B. (2010). Dynamic relief-demand management for emergency logistics operations under large-scale disasters, *Transportation Research Part E: Logistics and Transportation Review*, 46(1), 1-17
7. Yi, W. and L. Odamar, A dynamic logistics coordination model for evacuation and support in disaster response activities. European Journal of Operational Research, Vol. 179, No. 3, 2007, pp. 1177-1193.
8. Ortuño, M. T., Cristóbal, P., Ferrer, J. M., Martín-Campo, F. J., Muñoz, S., Tirado, G., & Vitoriano, B. (2013). Decision Aid Models and Systems for Humanitarian Logistics. A Survey. In Decision Aid Models for Disaster Management and Emergencies (pp. 17-44). Atlantis Press.
9. Ozdamar, L. and Ertem, M.A., Models, Solutions and Enabling Technologies in Humanitarian Logistics, European Journal of Operational Research (2014)
10. Huang, A., Ma, A., Schmidt, S., Xu, N., Zhang, B., Meineke, L. and Shi, Z.(E.), Chan, J. and Dolinskaya, I. (2013). Integration of Real Time Data in Urban Search and Rescue. Center for the Commercialization of Innovative Transportation Technology, Transportation Center, Northwestern University. http://www.ccitt.northwestern.edu/documents/Integration_of_real_time_data_in_urban _search_ and_rescue.pdf
11. International Federation of Red Cross and Red Crescent Societies (2013). World Disasters Report. http://www.ifrc.org/en/publications-and-reports/world-disasters-report/world-disasters-report-2013/
12. United Nation Foundation, Disaster Relief 2.0: The Future of Information Sharing in Humanitarian Emergency. http://www.unfoundation.org/what-we-do/legacy-of-impact/technology/disaster-report.html .
13. Typhoon Haiyan "Typhoon Haiyan Wiki Page," Retrieved on 28 January 2014 from
14. The National Oceanic and Atmospheric Administration "NWS Director Uccellini speaks on CNN about Typhoon Haiyan and Climate Change," Retrieved on 28 January 2014 from <http://www.nws.noaa.gov/com/weatherreadynation/news/131121_louis.html>
15. ReliefWeb. NDRRMC Updates re Effects of TY YOLANDA (HAIYAN) (as of 17 Apr 2014). National Disaster Risk Reduction and Management Council. http://reliefweb.int/sites/reliefweb.int/files/resources/Update%20Effects%20TY%20YOLANDA%2017 %20April%202014.pdf Accessed 28 May 2014
16. Office of Coordination of Humanitarian Affairs. Philippines: Super Typhoon Haiyan makes landfall, UN and humanitarian community on high-alert. http://www.unocha.org/roap/top-stories/philippines-super-typhoon-haiyan-makes-landfall-un-and-humanitarian-community-high-alert. Accessed 28 January 2014.
17. Digital Humanitarian Network. Super Typhoon Haiyan. http://digitalhumanitarians.com/content/super-typhoon-yolanda Accessed 7 April 2014.
18. Humanitarian Response. Humanitarian Response Philippines. https://philippines.humanitarianresponse.info/search/ . Accessed 11 January 2014.
19. Open Street Map. Index of Haiyan. http://labs.geofabrik.de/haiyan/ . Accessed 11 January 2014
20. United Nations Office of Coordination of Humanitarian Affairs. Who We Are. http://www.unocha.org/about-us/who-we-are . Accessed 28 January 2014
21. United Nations Office of Coordination of Humanitarian Affairs What is UNDAC?. http://www.unocha.org/what-we-do/coordination-tools/undac/overview . Accessed28 January 2014.
22. United Nations Office of Coordination of Humanitarian Affairs. Information Management. <http://www.unocha.org/what-we-do/information-management/overview . Accessed 28 January 2014.
23. CCCM Cluster. Typhoon Haiyan (Yolanda) Consolidated Cluster Briefs, 10 March 2014. http://reliefweb.int/report/philippines/typhoon-haiyan-yolanda-consolidated-cluster-br
24. Butler, Declan. Crowdsourcing goes mainstream in typhoon response. http://www.nature.com/news/crowdsourcing-goes-mainstream-in-typhoon-response-1.14186> Accessed 18 January 2014.
25. Official Gazette. NDRRMC report on actions taken by the government in the wake of typhoon Yolanda (November 11, 2013). http://www.gov.ph/2013/11/15/ndrrmc-report-on-actions-taken-by-the-government-in-the-wake-of-typhoon-yolanda-november-15-2013/
26. Official Gazette. Infographic: Timeline of government actions in rsponse to Typhoon Haiyan. http://www.gov.ph/2014/04/24/infographic-timeline-of-government-actions-in-response-to-typhoon-haiyan/
27. Open Street Map. Humanitarian OSM Team. http://wiki.openstreetmap.org/wiki/Humanitarian_ OSM_Team
28. Open Street Map. Typhoon Haiyan. http://wiki.osm.org/wiki/Typhoon_Haiyan
29. Liu, N. and Ye, Y. (2014). Humanitarian logistics planning for natural disaster response with Bayesian information updates. *Journal of Industrial and Management Optimization* 10(3), 665-689
30. Bogazici University, Earthquake Eng. Dept., 2002. Earthquake risk analysis of Istanbul metropolitan area (in Turkish). Bogazici Universitesi Kandilli Rasathanesi ve Deprem Arastirma Enstitusu, Istanbul, Turkey.
31. US Coastal Guard. Situational Awareness https://www.uscg.mil/auxiliary/training/tct/chap5.pdf Accessed 23 November 2014.
32. United Nations Office for Coordination of Humanitarian Affairs (2002). Symposium on Best Practices in Humanitarian Information Exchange. Palais des Nations Geneva, Switzerland.

33. Haggarty, A., Naidoo, S. (2008). Global Symposium+5 Final Report, Information for Humanitarian Action. UNOCHA Global Symposium Secretariat. Palais des Nations 1211 Geneva (2008)
34. Office of Coordination of Humanitarian Affairs (2013). Humanitarian Data Exchange. http://docs.hdx.rwlabs.org/wp-content/uploads/HDX-Project-Document-abridged.pdf
35. Office of Coordination of Humanitarian Affairs. Developing Humanitarian data Standards: an introduction and plan for 2014 http://docs.hdx.rwlabs.org/wp-content/uploads/HXL_Paper-forsite.pdf
36. http://docs.hdx.rwlabs.org/wp-content/uploads/HDX_Quality_Assurance_Framework_Draft.pdf
37. Management of a Crisis (MOAC) Vocabulary Specification. http://observedchange.com/moac/ns/
38. World Food Programme http://www.wfp.org/about
39. Logistics Cluster http://www.logcluster.org/logistics-cluster
40. http://www.wfp.org/fais/
41. https://www.fema.gov/pdf/about/pub1.pdf
42. http://www.fema.gov/data-feeds
43. http://www.fema.gov/femas-international-programs-activities
44. http://www.humanitarianinfo.org/iasc/pageloader.aspx?page=content-about-default
45. http://www.humanitarianinfo.org/iasc/downloaddoc.aspx?docID=5576type=pdf
46. Cordeiro, K.F.,Campos, M.L.M., Borges, M.R.S. (2014). Adaptive Integration of Information Supporting Decision Making: A Case on Humanitarian Logistic. Proceedings from: *11h International ISCRAM Conference*. USA.
47. Howden, M. (2009), "How humanitarian logistics information systems can improve humanitarian supply chains: a view from the field", Presented at *6th International ISCRAM Conference*. Gothenburg, Sweden.
48. Scott, N. and Batchelor, S. (2013) Real Time Monitoring in Disasters. *IDS Bulletin*, 44(2), 122-134.
49. Li, J., Li, Q., Liu, C., Ullah Khan, S., and Ghani, N. (2014). Community-Based Collaborative Information System for Emergency Management. *Computers and Operations Research*, 42, 116-124
50. Dorasamy, M., Ramen, M., and Kaliannan, M. (2013). Knowledge management systems in support of disasters management: A two decade review. *Technological Forecasting and Social Change*, 80(9), 1834-1853.
51. Palen, L., Anderson, K.M., Mark, G., Martin, J., Sicker, D., Palmer, M., and Grunwald, D. (2010). A vision for technology-mediated support for public participation and assistance in mass emergencies and disasters. Proceedings of *ACM-BCS Visions of Computer Science 2010*.
52. Munro, R. (2013). Crowdsourcing and the crisis-affected community. *Info Retrieval*, 16(2), 210-266.
53. Imran, M., Elbassuoni, S. M., Castillo, C., Diaz, F. and Meier, P. (2013). Extracting information nuggets from disaster-related messages in social media. Proceedings from: *ISCRAM 2013*, Baden-Baden, Germany.
54. Manso, M. and Manso, B. The Role of Social Media in Crisis: A European Holistic Approach to the Adoption of Online and Mobile Communications in Crisis Response and Search and Rescue. Proceedings from: *17th ICCRTS: "Operationalizing C2 Agility" International Command and Control Research and Technology Symposium*. http://www.dodccrp.org/events/17th_iccrts_2012/papers/007.pdf
55. Velev, D. and Zlateva, P (2012). Use of social media in natural disaster management. *International Proceedings of Economic Development and Research*, 39, 41-45.
56. Ortmann, J., Limbu, M., Wang, D., Kauppinen, T. (2011). Crowdsourcing Linked Open Data for Disaster Management. Proceedings of *Terra Cognita, the 10th International Semantic Web Conference*, Germany, 11-22
57. Hester, V., Shaw, A., and Biewald, L. (2010). Scalable crisis relief: Crowdsourced SMS translation and categorization with Mission 4636. Proceedings from *First ACM Symposium on Computing for Development*, p. 15.
58. Sarcevic, A, Palen, L, White, J, Starbird, K, Bagdouri, M and Anderson, K. (2012). "Beacons of Hope" in Decentralized Coordination: Learning from On-the-Ground Medical Twitterers During the 2010 Haiti Earthquake. Proceedings from *CSCW 2012*.
59. Limbu, M.K. (2012) Integration of crowdsourced information with traditional crises and disaster management information using linked data. ISEGI - MSc Dissertations Geospatial Technologies (Erasmus-Mundus). http://hdl.handle.net/10362/8305
60. Purohit, H., Castillo, C., Diaz, F., Sheth, A.P, Meier, P. (2014). Emergency-relief coordination on social media: Automatically matching resource requests and offers. *First Monday*, 19(1)
61. Ashktorab, Z., Brown, C., Nandi, M., and Culotta, A. (2014). Tweedr: Mining twitter to inform disaster response. Proceedings from: *11th International ISCRAM Conference*. USA.
62. Imran, M., Elbassuoni, S., Castillo, C., Diaz, F. and Meier, P. (2013). Practical Extraction of Disasterrelevant Information from Social Media. Proceedings from: *WWW 2013*, Rio de Janeiro, Brazil.
63. Tatham, P., Spens, K. (2011). Towards a humanitarian logistics knowledge management system. *Disaster Prevention and Management*, 20(1), 6-26
64. Humanitarian Response. Governmental Contact List - Region VI (Western Visayas). https://philippines.humanitarianresponse.info/document/government-contact-list-region-vi-western-visayas
65. World Bank Logistics Performance Index (LPI). http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/TRADE/0,,contentMDK:23188613 pagePK:210058 piPK:210062 theSitePK:239071,00.html
66. Tatham, P., L'Hermitte, C., Spens, K., and KovÃącs, G. (2013). Humanitarian logistics: development of an improved classification framework. Proceedings from: *11th ANZAM Operations, Supply Chain and Services Management Symposium*. Brisbane, Australia.
67. Teran, Javier (2014). Measuring the quality of humanitarian data: An emerging framework. http://docs.hdx.rwlabs.org/measuring-the-quality-of-humanitarian-data-an-emerging-framework/#comments
68. L'Hermitte, C., Bowles, M., Tatham, P.H. (2013). A new classification model of disasters based on their logistics implications. Proceedings from: *11th ANZAM Operations, Supply Chain and Services Management Symposium*. Brisbane, Australia.
69. Galton, A. and Worboys, M. (2011). An ontology of information for emergency management. Presented at *International Conference on Information Systems for Crisis Management and Response (ISCRAM)*. Lisbon, Portugal.
70. Jahre, M., Jensen L. (2010). Coordination in humanitarian logistics through clusters. *International Journal of Physical Distribution & Logistics Management*, 40(8/9), 657-674.
71. Palatino, Mong. Lessons from the Haiyan Typhoon Tragedy. *The Diplomat*.

72. World Health Organization (2013). Foreign Medical Teams in Tacloban City. http://www.humanitarianresponse.info/operations/philippines/infographic/foreign-medical-teams-tacloban-city . Accessed on 28 November 2013.
73. MapAction (2013). MA029 Philippines Typhoon Haiyan (Yolanda) 3W Oveview (21-Nov-2013). http://www.humanitarianresponse.info/operations/philippines/infographic/ma029-philippines-typhoon-haiyan-yolanda-3w-oveview-21-nov-2013 . Accessed on 28 November 2013.
74. Protection Cluster. Protection Cluster Displacement and 3W Map. http://www.humanitarianresponse.info/operations/philippines/infographic/protection-cluster-displacement-and-3w-map . Accessed on 28 November 2013.
75. Westrope, C., Banick, R., and Levine, M. (2014). Groundtruthing OpenStreetMap building damage assessment. *Procedia engineering*, 78, 29-39.
76. Geodata Converter. Vector. http://converter.mygeodata.eu/vector . Accessed 15 February 2014.
77. Vaillancourt, A 2013, 'Government decentralization and disaster impact, an exploratory study'. http://www.buildresilience.org/2013/proceedings/files/papers/352.pdf Accessed 12 June 2014.
78. Haavisto, I 2014, Performance in Humanitarian Supply Chains. Ekonomi och samhÃd'lle / Economics and Society, no. 275, Hanken School of Economics, Helsinki.
79. Ergun, O., J. L. HeierStamm, P. Keskinocak, J. L. Swann. 2010. Waffle House Restaurants hurricane response: A case study. International Journal of Production Economics 126(1) 111"ï£¡ 120.
80. MapAction. MapAction Home Page. http://www.mapaction.org/ . Accessed 9 January 2014.
81. Liu, S., Brewster, C. and Shaw, D. (2013). Ontologies for Crisis Management: A Review of State of the Art in Ontology Design and Usability. Proceedings from *10th International ISCRAM Conference*. Germany.
82. Gray, B. Crisis Mapping and its use in Disaster Management. Australian Fire Authorities Council. http://www.afac.com.au/downloaddoc.aspx?q=5b224dca-0f92-4ecb-acb0-19c1eb39cc5e. Accessed 15 June 2014.
83. Mansourian, A., Rajabifard, A., Valadan Zoej, M.J. and Williamson, I. (2006). Using SDI and web-based system to facilitate disaster management. *Computers* & *Geosciences*, 32(3), 303-315.
84. Digital Humanitarian Network. Purpose. http://digitalhumanitarians.com/about . Accessed 15 Feb 2014.
85. World Food Programme. WFP Logistics. http://www.wfp.org/logistics . Accessed 28 January 2014.
86. United Nations Office of Coordination of Humanitarian Affairs. Coordination Tools: Needs Assessment. <http://www.unocha.org/what-we-do/coordination-tools/needs-assessment . Accessed 28 January 2014.
87. United Nations Office of Coordination of Humanitarian Affairs. OSOCC & RDC. http://www.unocha.org/what-we-do/coordination-tools/osocc-rdc/overview . Accessed 28 January 2014.
88. Office of Coordination of Humanitarian Affairs (2013). World Humanitarian Data and Trends 2013. https://docs.unocha.org/sites/dms/Documents/WHDT_2013%20WEB.pdf
89. United Nations Population Fund (2010). Guidelines on Data Issues in Humanitarian Crises. http://www.unfpa.org/webdav/site/global/shared/documents/publications/2010/guidlines_dataissues.pdf
90. Karunakara, U. (2013). Data Sharing in a Humanitarian Organization: The Experience of Medecins Sans Frontieres. *PLoS Medicine*, 10(12).
91. United Nations Office of Coordination of Humanitarian Affairs. Situation Report No.30. http://reliefweb.int/sites/reliefweb.int/files/resources/OCHAPhilippinesTyphoonHaiyan SitrepNo 30_06January2014.pdf . Accessed 18 January 2014
92. Logistics Cluster. Logistics Cluster Operations Philippines Page. http://www.logcluster.org/ops/phl13a . Accessed 11 January 2014. f
93. Office of Evaluation, WFP; UNICEF Evaluation Office; and the Netherlands, Ministry of Foreign Affairs, Policy and Operations Evaluation Department. Joint Evaluation of the Global Logistics Cluster. <http://documents.wfp.org/stellent/groups/public/documents/reports/wfp251775.pdf>.
94. LogIK. Logistics Information about In-Kind Relief Detailed Report. http://logik.unocha.org/SitePages/DetailedReportAdmin.aspx?soid=2,3 . Accessed 11 January 2014.
95. LogIK. About LogIK. http://logik.unocha.org/SitePages/about.aspx . Accessed 11 January 2014
96. Open Street Map. Some editing stats from the Typhoon Haiyan response. http://hot.openstreetmap.org/updates/2014-01-14-some-editing-stats-from-the-typhoon-haiyan-response . Accessed 15 February 2014.
97. MapAction. Deployment Details. http://www.mapaction.org/deployments/depldetail/224.html . Accessed 15 February 2014.
98. MapAction,. Typhoon Haiyan Update. http://www.mapaction.org/more-news/385-typhoon-haiyan-update.html . Accessed 15 February 2014.
99. Copernicus Emergency Management System. EMSR058: Typhoon in Philippines. http://emergency.copernicus.eu/mapping/list-of-components/EMSR058 . Accessed 18 November 2014.
100. Copernicus Emergency Management System. What is Copernicus. http://emergency.copernicus.eu/mapping/ems/what-copernicus . Accessed 18 January 2014.
101. Copernicus Emergency Management System. EMS - Mapping Service. http://emergency.copernicus.eu/mapping/ems/ems-mapping-service . Accessed 18 Jan 2014.
102. ESRI. Cyclone Haiyan Impact (MapServer). http://fema-dev.esri.com/arcgis/rest/services/DisasterResponse/Cyclone_Haiyan_Impact/MapServer . Accessed 25 November 2014.
103. All Partners Access Network. Typhoon Haiyan Response 2013. https://community.apan.org/typhoon-haiyan/p/map.aspx . Accessed 6 Dec. 2013.
104. Digital Humanitarian Network. Esri Disaster Response Program. http://digitalhumanitarians.com/content/esri-disaster-response-program . Accessed 8 April 2014.
105. ESRI. Typhoon Haiyan/Yolanda Maps. http://www.esri.com/services/disaster-response. Accessed 8 April 2014.
106. ESRI. Typhoon Haiyan/Yolanda Maps. http://www.esri.com/services/disaster-response/hurricanes/typhoon-hayian-yolanda-maps . Accessed 15 February 2014
107. UN Institute for Training and Research. UNOSAT. http://www.unitar.org/unosat/ . Accessed 20 November 2013
108. VISOV. Download Reports. https://haiyan.crowdmap.com/reports/download . Accessed 18 January 2014.
109. VISOV. Volontaires Internationaux en Soutien aux Opérations Virtuelles. http://visov.org. Accessed 19 January 2014.
110. VISOV, "About Us," Retrieved on 11 Jan. 2014 <https://haiyan.crowdmap.com/page/index/1>

111. National Disaster Risk Reduction and Management Council. NDRRMC Home Page. http://www.ndrrmc.gov.ph. Accessed 11 January 2014.
112. Department of Social Welfare and Development. Reports and Updates. http://disaster.dswd.gov.ph/reports-and-updates/ . Accessed 11 January 2014.
113. Humanitarian Response. About Humanitarian Response. https://philippines.humanitarianresponse.info/about . Accessed 18 January 2014.
114. Humanitarian Response. Humanitarian Response COD-FOD Registry Philippines. https://cod.humanitarianresponse.info/search/field_country_region/164?search_api_vi ews_fulltext= . Accessed 11 January 2014.
115. ReliefWeb. ReliefWeb Home Page. http://Reliefweb.int . Accessed 09 January 2014.
116. ReliefWeb. ReliefWeb About Page. http://reliefweb.int/about . Accessed 18 January 2014.
117. All Partners Access Network. Typhoon Haiyan Community. https://community.apan.org/typhoon-haiyan/b/updates/archive/2013/11/09/typhoon-haiyan-community.aspx . Accessed 18 January 2014
118. Red Cross. Typhoon Haiyan Mapfolio. http://americanredcross.github.io/haiyan_mapfolio/ . Accessed 3 December 2014.
119. Google Crisis Maps. Typhoon Yolanda Relief Map. http://google.org/crisismap/2013-yolanda . Accessed 11 January 2014.
120. Google Crisis Response. Tools For Responders. https://www.google.org/crisisresponse/ . Accessed 11 January 2014.

Appendix

Data Sources Classification

The information posted at the outlets, generally in their "about us" and data pages are analyzed based on their relevance to humanitarian logistics models during the timeline of the study, between November 10, 2013 and March 31, 2014. In Section 3.3, we identify two types of information outlets: information aggregation sites and organization or sector specific data repositories (original, in other words, primary outlets). However, this dichotomy may not necessarily be entirely distinct, as some organization-focused sources will also include some data collected from other groups, For example, MapAction takes advantage of data from NDRRMC to generate some of the maps. Thus, one might expect to categorize MapAction as an aggregator. Also, there might be more information outlets providing information after Typhoon Haiyan. This classification includes analysis of a subset of available resources. Table 1 comprises a subset of the proposed classification categories. Some of them require further analysis. For example, at this point, we are unable to fully judge retention time of each source since, as of March 2014, this is still an active crisis.

Original Information Outlets

Logistics Cluster [92]
Logistics Cluster, created by OCHA [93], aids the cooperation of groups of humanitarian organizations. World Food Programme (WFP) is the lead agency as appointed by the IASC [92][93]. Common types of datasets are maps, meeting minutes, and situation updates. Logistics Cluster maintains maps that detail infrastructure data such as operations access constraints and general operations. The meeting minutes from the Coordination - Roads Transport - Sea and Rivers Transport Group include information about the road conditions. These almost daily meeting notes start from November 11, 2013. The situation updates also provide information about various types of transportation channel availability including the overland transport. These updates start on November 14, 2013, with limited data about the road conditions. Most of Logistics Cluster's files found during the case study comprise of portable data formats (PDF).

LogIK [94]
LogIK, or Logistics Information about In-Kind Relief, is a global online database maintained by OCHA. LogIK provides detailed reports in PDF or XLS format of supply operations, with three categories of data: relief items, transports and contributions. The database reflects reported international/ regional humanitarian contributions of relief items. Within these categories, LogIK offers information such as supplier information, destination, and quantity [95]. Specifically, the relief items section includes data about donated item types (e.g., blankets, tents and emergency kits), quantity, senders and other. The contributions section provides the decision date and dollar value of the contributions. The transport section contains information about vehicles provided from different organizations by air, road and sea. This information source may be of higher reliability because its source data originates from donors affiliated with the United Nations Office (UN). The information in LogIK is updated daily.

HOT [27]
The Humanitarian OpenStreetMap Team (HOT) is a volunteers based community formed within the larger OpenStreetMap community that has emerged as a pivotal provider and platform for data in humanitarian operations by providing open source data [27]. The data placed into OSM by volunteers continues to increase its scope and accuracy with rising numbers of users verifying information in more locations [12]. OSM furthermore contributes to the large growth in information in post-disaster operations by frequently updating geographic data, sometimes every minute [28]. OpenStreetMap globe data takes several gigabytes, so specific repositories exist for disasters such as Haiyan. While the "history" feature [19] of the OSM helps to see the previous actions, space limitations tend to prompt these repositories to update at longer intervals and not retain many previous updates . The initial OSM updates for Typhoon Haiyan dates back to November 7, 2013 since the HOT team was called by OCHA to start mapping the region a day before the typhoon touchdown [24]. The basic street maps of the cities of Port-Au-Prince and Carrefour provided by OpenStreetMap in about 48 hours following the previous crises were claimed to be the best available maps [96].

MapAction [80]
MapAction is a non-governmental organization that produces maps for the humanitarian crisis. From November 13, 2013 to January 17, 2014, they provided maps in JPEG and PDF formats of affected areas, along with the information on the populations, road

Table 1 Classification of Information Data Files Collected by Our Team (see Section 3.3 for details on classification categories and Appendix for description of sources)

Source Name	Outlet Type	Primary Purpose Posted Related to HL Models	Logistical Content	Main Downloaded File Type(s)	Update Frequency Observed in Timeline	Update Type Observed	Establishment	Update Timeline	International / Local
APAN	aggregator	user-uploaded files	all	mixed	varies	new	disaster-specific	11/12/13 - *present**	international
Copernicus EMS	primary	initial post disaster assessment	infrastructure	SHP	daily	new	multiple disasters	11/12/13 - 11/18/13	international
DSWD	primary	evacuation	supply, demand	PDF	every few days	new	multiple disasters	11/09/13 - 12/12/13	local
Google Crisis Maps	aggregator	facility placement and damage assessment	infrastructure	KML	unknown	overwrite	multiple disasters	11/08/13 - 11/26/13	international
Humanitarian Response	aggregator	satellite imagery	all	mixed	daily	new	multiple disasters	11/12/13 - *present**	international
IFRC	aggregator	displacement	all	PDF	daily	new	disaster-specific	11/10/13 - *present**	international
Logistics Cluster	primary	coordination and resource aggregation	infrastructure	PDF	almost daily	new	multiple disasters	11/12/13 - *present**	international
LogIK	primary	relief aid	supply	XLSX	daily	overwrite	multiple disasters	unknown	international
MapAction	primary	disaster response maps	all	PDF	daily	new	multiple disasters	11/12/13 - *present**	international
NDRRMC	primary	facility Operations and Damage Assessment	All	PDF	daily	new	multiple disasters	11/08/13 - *present**	local
OSM (repository)	primary	user-mapped damage, damaged infrastructure	infrastructure	SHP, OSM	daily	overwrite	multiple disasters	11/08/13 - *present**	international
Reliefweb	aggregator	aggregates various sources	all	mixed	daily	new	multiple disasters	11/08/13 - *present**	international
UNITAR - UNOSAT	primary	initial post disaster assessment	infrastructure	PDF	daily	new	multiple disasters	11/12/13 - 11/20/13	international
VISOV	primary	user-mapped damage	infrastructure	CSV	daily	new	disaster-specific	11/09/13 - *present**	international

∗ Timeline of the reported team's actions is between November 10, 2013 and March 31, 2014. As of March 31, 2014, the updates are still being posted.

conditions, coordination (cluster activities by location), shelters and others. The main type of MapAction maps seem to be "Who, What, Where" maps that exhibit the locations of organizations. MapAction compiles data from several sources such as OCHA and NDRRMC for different cities, and each map is accompanied by a summary. A lag appears to exist between the report date and the update time; however, MapAction directs its users to Humanitarian Response Philippines portal for primary MapAction maps [**?**]. Similar to OSM, MapAction was also present in Philippines before the typhoon hit [98].

Copernicus EMS [99]

CEMS (Copernicus Emergency Management System), maintained by the European Commission, "monitors and forecasts the state of the environment on land, sea and in the atmosphere, in order to support climate change mitigation and adaptation strategies, the efficient management of emergency situations and the improvement of the security of every citizen" [100]. The website appears to present these maps, which seem to run between November 9, 2013 to November 18, 2013, in numerous formats and resolutions for over a year after the disaster. Moreover, while Copernicus EMS is claimed to have published some of the best pre- and post-event analysis images in the first 36 hours of the Haiyan's landfall [24] , the website appears to make only few updates publicly once the initial assessment occurs.

ESRI [102]

Environmental Systems Research Institute (ESRI) holds data from the US Government on infrastructural damage [103]. The ESRI Disaster Response Program (DRP) and entity with the organization, supports organizations responding to disaster. They provide "software, data coordination, technical support, and other GIS assistance to organizations" [104]. In this case study, similar to CEMS, these files appear to consist of initial damage assessments. They supported the Typhoon Haiyan response by providing an ESRI platform for publically license imagery after the event, and have support other disasters [105]. ESRI was one of the organizations that responded to OCHA call for volunteers as part of the DHN network. After Typhoon Haiyan, ESRI collaborated with the digital volunteer mapping groups such as Standby Taskforce and GISCorps to process social media reports and provide interactive maps [106]. The website also provides maps from other groups such as MapAction and OSM. The Haiyan maps start from November 8, 2013 and were last updated on November 25, 2013 (as of March 2014).

UNITAR - UNOSAT [107]

United Nations Institute for Training and Research's (UNITAR) Operational Satellite Applications Programme (UNOSAT) is a satellite program that provides "solutions to relief and development organizations within and outside the UN system to help make a difference in critical areas such as humanitarian relief" [107]. The satellite images appear to allow digital mapping volunteers to contribute to changing sources, such as OSM, and often remains available for several years. Daily maps illustrating brief overview of satellite-detected areas of destroyed and possibly damaged structures of different areas of Philippines are provided from November 11, 2013 to November 20, 2013. While the first few are presented only in PDF format, the rest are also offered in Shapefile and ESRI's geodatabase format.

VISOV [108]

The goal of Volontaires Internationaux en Soutien aux Opérations Virtuelles or International Volunteers in Support of Virtual Operations (VISOV) "is to help coordinate disaster responses with those of emergency organizations (formal or humanitarian) via digital spaces on which they organize and communicate" [109]. VISOV appears to openly share and maintain its datasets on the website, possibly due to its intention to "become a tool in the hands of local communities" [110]. These datasets contain relevant tweets and map tags to estimate the road damage and relief progression [110]. VISOV datasets in particular include information such as the type of damage, description of the damage, geographical location, and time of notification. The data are available in the comma-separated value (CSV) and keyhole markup language (KML) format from November 11, 2013 to December 3, 2013.

NDRRMC [111]

NDRRMC, a governmental agency of the Philippines, develops detailed situation reports used by many mapping efforts and other situational reports. These PDF reports include information about situation overview, causalities, affected population, damaged houses, status of roads and bridges, standees, prepositioned and deployed assess/resources, cost of assistance, cost of damages, status of lifelines (both power and network outage), and emergency management. The status of the roads and bridges demonstrates the damaged areas, declares if the roads are passable and adds remarks such as closing reasons or efforts made to make the roads passable. The level of detail includes even missing persons' names, as well as the coordination efforts (involvement of different governmental and international agencies and humanitarian groups). These reports were initiated immediately after the Haiyan on November 8, 2013. NDRRMC retains the situational reports during the recovery operations and appears to archive a large number of files.

DSWD [112]

The Department of Social Welfare and Development (DSWD) appears to play a similar role to NDRRMC. However, it seems to focus on breaking down the information on citizens by geographic regions, as well as statuses within each region such as the number of families in each evacuation center. For the Typhoon Haiyan, DSWD frequently publishes effect, service and intervention reports from November 8, 2013 to December 12, 2013. Viewing previous disasters suggests that DSWD also retains its reports for several months after the onset of the disaster.

Information Aggregation Outlets

Humanitarian Response [18]

Humanitarian Response, maintained by OCHA, "aims to be the central website for Information Management tools and services" [113]. It appears to compile files from OCHA sectors (Logistics Cluster, Education Cluster, Protection Cluster, etc.), and other groups such as the Canadian Red Cross, Logistics Cluster, MapAction and OSM. The Humanitarian Response website possesses a large number of files, retaining information from several past operations. The outlet provides numerous file filters such as content and data source, and within each filter, multiple items may be selected [113]. Humanitarian Response also maintains a registry of common operational data sets and fundamental operational data sets (COD - FOD registry) that often contains files with numerical data, which it claims "should represent the best-available data sets for each theme" [114]. The relevant data starts from as early as the moment the Typhoon Haiyan hit, and new information is still being uploaded months after the event.

ReliefWeb [115]

As with Humanitarian Response, OCHA maintains the ReliefWeb website. ReliefWeb appears to differ from Humanitarian Response in that it provides files, from situation reports to maps, from a broad range of sources and topics, not focusing on information management to the extend that Humanitarian Response does. ReliefWeb may be effective for identifying primary sources, since it “collects, updates and analyzes from more than 4,000 global information sources” [116]. Alternatively, ReliefWeb may help narrow which sources’ files do not need to be captured right away since it appears to contain most of the files from each source and retains them long after relief operations. The OCHA sourced information about Typhoon Haiyan is directly linked to ReliefWeb website on the OCHA website. ReliefWeb page for Typhoon Haiyan was activated on November 8, 2013 and different updates are still being uploaded, as of March 2014.

APAN [103]

All Partners Access Network (APAN) functions similarly to ReliefWeb and Humanitarian Response, but differs mainly in that users upload the files themselves and that the specific page for Typhoon Haiyan is reactionary [117]. User uploaded files allow for identification of reactionary sources that may be overlooked in the expansive collection of ReliefWeb and Humanitarian Response sources. However, user uploading tends to lack consistency in uploading files from any given primary source, so using APAN as a data retrieval site may be problematic. In contrast, users may sometimes upload files not on a given website but derived from non-public datasets, e.g., insurance industry datasets. APAN amalgamates maps, briefs, reports from a variety of different organizations, agencies and groups from November 10, 2013 to January 7, 2014. The community for Typhoon Haiyan provides a link to an ESRI map.

Red Crescent Societies (BRC, ARC) [118]

The Red Crescent Societies do not appear to put out files as an overarching system of organizations; rather some individual Red Cross societies may choose to do so on their own. In this case, a collaboration between the American Red Cross and British Red Cross [118] makes available numerous maps throughout the disaster recovery efforts using various data sources. Since the map files specify what data sources each map employs, they may be used to locate the primary sources that contain the desired raw data. Moreover, the maps seem to specify the exact file, e.g., report number, from the source, allowing for direct retrieval of specific data. Red Cross also provides reports about damage assessments, affected people, shelter, etc. They collect information from a variety of resources, such as OSM, UNITAR-UNOSAT, and ReliefWeb.

Google Crisis Maps [119]

Google Crisis Maps, one of the tools of Google Crises Response Group [120], crowdsources data not only within its self-produced facility locations files [119], but also by providing options to access files from sources such as Waze, a traffic mapping application, and CNES/Astrium, which provides satellite imagery. In particular, the self-produced map from Google Crisis Maps plays a role as infrastructure data. However, the facilities that Google Crisis Maps displays appear to remain relatively constant at each map, so frequent downloads may not be necessary depending on the goals. The map shows damaged areas, their severity, evacuation centers, and relief drop zone areas. When color-coding the damaged areas, the map shows the data as aggregated chunks. However, it is not always clear if this means that the roads to those areas or the roads within that area are closed or not; and more detailed explanation about classification of damages might be useful.