Humanitarian Technologies and Big Data

A critical approach for operational effectiveness

THE BIG PICTURE

A serious scholarly approach is needed to explore the opportunities and challenges of harnessing big data generated by mobile technologies, social media, commercial remote sensing, and unmanned aerial vehicles (UAVs) for humanitarian response and disaster management. Such a research agenda will not only augment the operational capacity of the humanitarian community, but—and this is critical—will also contribute to an understanding of what these technologies do to social and power relationships. As Sandvik et.al point out, technology itself cannot be thought of in normative terms, nor can it be conceived of as a neutral vessel: it reflects a mutually constitutive relationship with society.[1]

Opportunities and progress

The proliferation of access to low and high grade technologies has generated innovative, timely, and cost efficient ways of capturing and making sense of actionable intelligence in crisis affected settings.

The Satellite Sentinel Project showed how near-real time satellite imagery could provide both a conflict early warning system and a living archive for the documentation of human rights violations in the border region of South Sudan.[2] The Nethope / Standby Task Force Ebola Phase II deployment illustrated how a network of remote-based digital volunteers could provide the World Health Organization with its first comprehensive map of Ebola Treatment Facilities in West Africa during the height of the crisis.[3] And Micro-Mappers, Humanitarian Open Street Map, UAviators, Kathmandu Living Labs, and the Standby Task Force—all digital humanitarian, volunteer and technical communities (V&TCs)—exemplified how traditional and new media, satellite imagery, and high-resolution video footage could be aggregated, interrogated, and analyzed to provide

[1] Kristin Begtora Sandvik, Maria Gabrielsen Jumbert, John Karlsrud and Mareile Kaufmann, "Humanitarian Technology: a Critical Research Agenda," *International Review of the Red Cross*, 2014.
[2] See Satellite Sentinel Project "Making the World a Witness: Report on the Pilot Phase December 2010 – June 2012," (Harvard Humanitarian Initiative: July 2012). http://www.satsentinel.org/report/making-world-witness-report-pilot-phase-report
[3] See Joseph Guay, "Digital Humanitarians are Improving the Ebola Response," (ALNAP: January 2015). http://www.alnap.org/blog/121

situational awareness on infrastructure damage and the locations and needs of vulnerable populations in real time after a magnitude 7.8 earthquake.[4]

Indeed, in an increasingly networked world, where innovations in—and access to—technology are growing at an exponential rate, the possibilities seem endless. Already, the International Federation of the Red Cross and Red Crescent Societies (IFRC) have launched a Tech4Resilience Global Dialogue on Emerging Technology for Emerging Needs. According to IFRC, augmented reality software (software that adds layers of computer-generated data through internet-connected devices) might be used for "crowdsourcing and visualizing community resources," in disaster settings." The Dialogue is also engaging with applications of biometric scanners and data systems for registration of vulnerable communities, smart home technology for remote monitoring of build environments, and wearable technologies that monitor health conditions and mobility of vulnerable populations, augmenting real-time communication and navigation for early warning and aid search and rescue in situations of complex crises.

Challenges and emergent problems

But while new humanitarian technologies, actors, and ways of doing business are setting the bar for actionable intelligence during mass atrocities, pandemics, and natural disasters, the sector as a whole lacks the necessary frameworks and protocols for accurately, appropriately, and safely harnessing these tools and methods. Micro-tasking open-source data and social media by volunteer organizations can provide timely and cost efficient data collection and analysis but requires the training and professionalization of traditional humanitarian organizations in order to achieve accurate and reliable results. The use of satellite imagery and UAVs by non-state actors brings military-grade intelligence, but without the diplomatic tools and legal frameworks necessary to deal with issues of potential violations of sovereign airspace. Using algorithms to mine meta-data from Twitter and personal call detail records bring real time holistic opportunities but also privacy issues and the potential to empower some segments of society at

[4] See Joseph Guay, "Saving Nepal: The Information Revolution," (Brookings TechTank: May 2015). http://www.brookings.edu/blogs/techtank/posts/2015/05/1-nepal-information-revolution.

the expense of others. Policy implications of using smart sensors and the internet of things (IoT) for development are still being hashed out.[5]

How are we to navigate these emergent issues? If mobile and high-end technologies are to contribute to information management in crisis contexts, organizations need to deal with the following:

(1) coordination and collective action problems
(2) standardization and integration of divergent, interoperable data streams
(3) privacy, security, and ethical challenges
(4) credibility, accuracy, and verification gaps
(5) matters of sovereignty, legitimacy, participation, and relationships.

The following cases exemplify a broad range of challenges and opportunities to leverage new technologies and big data in humanitarian action.

[5] See CISCO and ITU, "Harnessing the Internet of Things for Global Development" (2016): http://www.itu.int/en/action/broadband/Documents/Harnessing-IoT-Global-Development.pdf. See lessons learned with SENSORS: http://www.ictworks.org/2015/11/23/sensors-for-merl-what-works-what-does-not-what-have-we-learned/.

CASE STUDIES AND ONGOING PROJECTS

Nepal Earthquake

Making sense of the information explosion during natural disasters: standardization and integration of divergent data streams for decision-making.

In the networked age, scarcity of information is no longer a concern.[6] Just the opposite; organizations need to figure out how to make sense of *too much data* made accessible by mobile technologies, social media, and access to visual information through more affordable commercial remote sensing technologies and UAVs. Such work is like finding the proverbial needle in the haystack, and requires establishing mechanisms to *limit* information, clean and filter data generated through divergent sources, and share that information horizontally and in near-real time.

The Nepal earthquake response is one example of an explosion (and convergence) of divergent data and the need for critically assessing information products for decision-making.[7]

At the outset of the earthquake, Digital Globe and Google's Skybox donated satellite imagery to the humanitarian community, free of charge. Facebook and Google launched Safety Check and Person Finder apps to locate lost loved ones. Micro Mappers and Humanitarian Open Street Map (HOT) were activated (with official United Nations support). Artificial Intelligence for Disaster Response (AIDR) was used to process and filter the hundreds of thousands of Tweets pertaining to the #NepalEarthquake and #NepalEarthquareResponse, while Micro Mappers process tweets and images by the click of a mouse. The Standby Task Force was activated to support information management and geo-spatial mapping

[6] See "Disaster Relief 2.0," (Harvard Humanitarian Initiative: 2013).
[7] Guay, "Saving Nepal" 2015. For example,

of information resources for responders. A combination of government census data and imagery analysis has been uploaded to the Humanitarian Digital Exchange (HDX).

In short, a multitude of data were processed by an ecosystem of actors to produce a variety of information products germane to infrastructure damage or the location and needs of affected communities. In just one week after the crisis for instance, I counted over 100 data exchanges alone, making available static reports, PDFs, spreadsheets, and links to interactive maps and other exportable baseline layer data.

But what of the users of such information products? Humanitarian relief involves the complex interplay of governments, international organizations, NGOs, local organizations and affected communities themselves. How are these actors making sense of the information deluge coming out of crises contexts? What about affected communities themselves? In what ways have donated satellite imagery and UAV footage served to inform vulnerable and displaced populations? What ways can the information cycle be completed through context appropriate feedback loops? Current research suggests limited ways of knowing just how effective these new products were.

While the proliferation of communications tools, information platforms, and data exchange has indeed increased the potential for more temporal-spatial and category specific situational awareness—in near-real time—it has also led to an "innovation-action gap," as a diverse range of humanitarian actors are still struggling to digest, make sense of, manage, and use information for decision making in a variety of complex crises situations. Part of this has to do with the supply-driven nature of information being pumped out of crisis contexts. Supply-driven information products generate the following challenges for humanitarian actors:

(1) Too much or irrelevant information. For information to be a resource, is has to be aligned with and targeted to task. In the "information age," the increase in velocity, volume, and variety of data (big data) and the operational context of

humanitarian work (new actors, threats, contexts, and technologies) make decision making especially challenging.

(2) Those who produce information lack knowledge about the users of such information. Tools, platforms, and data exchanges are supplied at an ever increasing rate and in multiple varieties due to technological advancements that have reduced barriers for suppliers of information. But information products might not be aligned with user needs (a variety of demands).

(3) Information biases. Increasing the readability, search-ability, and timeliness of information cannot address the problem of political positioning that often emerges during dialogue and remains a barrier to the use of information for multi-organizational, multi-level, and multi-disciplinary collaboration between and among digital and traditional communities.

We need to design and test "user-driven" approaches, shifting from supply to a demand-based orientation toward information products in crisis situations, with the goal of creating a new kind of solution that not only provides information but also facilitates its use. To do this, we would enage with a diverse community of decision makers, collect evidence on challenges and opportunities, and develop a prototype for a practical tool to help first responders navigate an ecosystem of information for decision-making . When working with NGOs for example, we would collaborate with information management officers and response managers to see if and how such diverse information products are integrated (or not) into Emergency Management Systems. In our work with affected communities, we would first survey information infrastructure and socio-cultural contexts to imagine ways in which complex information can be harnessed to complete feedback loops to more effectively engage with affected communities in the information management cycle.

Such a tool will provide snapshots of activities to help actors understand the complexity of information available, aid users in navigating lean pathways toward effective application of information that avoids the challenges often associated with supply-driven information production, and will serve to evaluate the effectivenss of new information products to facilitate truly interoperable platforms.

Containing Ebola

Un-siloing data systems and bridging the "digital/physical divide" for coordinated response in public health emergencies.

Remote management in high-risk crisis situations is now possible due to advancements in ICTs, even in traditionally low-resource settings. Such an approach offers up new opportunities for multi-modal governance in areas of limited statehood.[8] But this practice also raises questions about distancing tendencies, as on-ground presence gives way to new "digital proximities," contributing to a re-allocation of risk away from the humanitarian organization toward local communities, and leading to an overestimation of the degree of beneficiary empowerment and participation through ICTs.[9] Facilitating shifting roles and new relations may require new rules for data governance to standardize information management procedures, integrate data collection systems, and encourage the right amount of competition in an environment that is traditionally characterized by duplication of efforts, silo-ed approaches, and competition between players.

The Ebola response in West Africa in 2014 is a prime example of this.[10] By October 2014, many months after the WHO declared an Ebola outbreak emergency in West Africa, in the international response was insufficient, uncoordinated, and unable to curb the spread of infection in the worst hit countries. Information management problems are considered to be a primary factor for the inadequate response leading up to the height of the crisis. At the time, facility-centric information (about the location and capacities of Ebola Treatment Units) was silo-ed within the NGOs that were in administrative charge of the facility. A Save the Children facility, for example, had little information (location, supplies, case load, capacity), about an Medicines Sans Frontieres

[8] See Steven Livingston and Joseph Guay, "Climate Change Induced Displacement: Leveraging Transnational Advocacy Networks (TANs) to Address Operational Gaps," *Brookings TechTank,* March 2015 here: http://www.brookings.edu/blogs/techtank/posts/2015/03/3-transnational-advocacy-networks-livingston-guay.
[9] Sandvik et. al 2014.
[10] See Joseph Guay, "Digital Humanitarians are Improving the Ebola Response," (ALNAP: January 2015). http://www.alnap.org/blog/121.

(MSF) facility close by, and vice versa. This negatively impacted vector monitoring and regional resource management.

Luckily, digital humanitarians stepped in to help, leveraging crowd-sourced and open-sourced methods and innovative mapping platforms. In just over three weeks time in October, the Standby Task Force deployed through Net Hope and in collaboration with UN OCHA's Humanitarian Data Exchange (HDX) to produce the most comprehensive and up-to-date snapshot of operational treatment facilities in the region. Before the SBTF/Nethope deployment facility-centric data (about the locations and attributes of treatment centers and health clinics) did not exist in a shared centralized database. However, because of the limits of remote and open-sourced data mining, the data generated by the crisis mapping partnership was incomplete, unverified, unstructured, and, upon publication, instantly out of date. Some health facilities did not indicate which NGOs were in administrative charge or include basic information about number of beds, PPE kits, staff or lab capacity. Check out this map here—lots of question marks.

What was needed (and ultimately not established) was a field enumeration team on the ground to augment the data using mobile survey tools to gain and update information, in real-time, about capacity attributes of the facilities. In December of 2014, I envisioned such information to be readily available to donors, coordinating bodies, NGOs, and even affected community members through a lightweight "data dashboard." We thus proposed such a solution with MIT Global Ideas Challenge and held talks with UNMEER.[11] This initiative, however, was tabled due to financial, security, and other bureaucratic reasons, including the difficulty in quantifying value-added and generating user by-in.

Moving forward, it will be crucial to articulate frameworks by which digital humanitarians can not only augment intelligence for field workers, but for field staff to contribute needed information for remote analysis. We need to pursue solutions for two-way engagement between remote-based digital volunteers and field staff (vertical integration) as well as for sharing across organizations and staekholders (horizontal integration).

[11] See "Ebola Treatment Facilities Data Dashboard" project pitch at MIT Ideas Global Challenge here: http://globalchallenge.mit.edu/problems/view/110.

4Mi

Using mobile survey tools and remote management to fill data gaps on mixed migration in transit countries in the Horn of Africa

For humanitarian organizations and host country governments information on the drivers, profile, conditions, and routes of transit for vulnerable communities on the move is crucial for mixed migration decision-making and policy. And yet there is a dearth of information on such indicators, especially in transit countries, which limits our capacity to anticipate movement trends or design context-appropriate interventions in countries of origin, transit, and destination.

Participatory research experience with data collection initiatives indicates that actionable intelligence in low-resource, high-risk settings is possible, although validity and ethical/security concerns remain particularly relevant.[12]

4Mi is a research initiative operating out of the Regional Mixed Migration Secretariat Nairobi office that looks to contribute timely and meaningful intelligence in an otherwise sparse data environment on mixed migration in the Horn of Africa Region. The initiative is an innovative, low cost, community-based approach to monitoring mixed migration flows, unlocking the potential to contribute timely, rich, and sensitive information about the drivers and trends of mixed migration, the condition of migrants, and revealing the clandestine nature of smuggler networks and routes. The team uses a mobile survey application and a network of local informants to collect data. After the Pilot Phase (July-2014 through end of 2015) the team has collected and analyzed over 800 survey reports from monitors in Egypt, Ethiopia, Kenya, South Africa, Turkey, and more.

However, due to a lack of ethically-informed and empirically-driven design (EBD)—an approach especially important when dealing with vulnerable populations in high-risk settings and when new tools and methods are built and

[12] See 4Mi Evaluation Concept Note, on file with author and Policy Lab.

used—a number of weak spots hinder the contribution of the initiative. Because an EBD was not used during the pilot stage, major questions of (1) validity and (2) the ethical and safe use of data have emerged, which if not addressed, could lead to credibility and liability problems at time when 4Mi is being pressured to publish results for policy-makers and scale up and scale out their methodology across contexts.

Such an initiative is representative of the challenges for this kind of work across organizations and sectors. Actors seeking to leverage mobile collection technologies and remote management of local informants need to build in an evidence-based design approach to consider accuracy and reliability issues related to elements of survey design, sampling methodology, the training of local informants, and validation and quality control. Also important will be to consider matters of privacy, ethics, and security regarding collecting information on vulnerable populations (migrants) and perpetrators of criminal activity (smugglers and potential human traffickers).[13]

Indeed, a major gap in mobile survey tools is the lack of minimal standards and professional ethics by which to guide the application of new ICTs for humanitarian action.[14] According to Signal Program for Human Security and Technology Director Nathaniel Raymond, "there exists little or no accepted operational doctrine, nor precedents for applying traditional humanitarian principles to ICT-supported operations."[15]

A dearth of doctrine is largely due to the way that humanitarian technological applications are born: advancements around new technologies and big data in the humanitarian sector have largely been driven by pilot projects and/or organically-generated collaborations and initiatives that emerge when new capacities and tools can meet contextual demands.[16] But it is difficult to wrestle with questions of

[13] See 4Mi Policy Lab Concept Note for details on validity and ethical/security-related challenges.
[14] See Nathaniel Raymond and Brittany Card, "Applying Humanitarian Principles to Current Uses of Information Communication Technologies: Gaps in Doctrine and Challenges to Practice," *Harvard Humanitarian Initiative Working Paper,* July 2015 here: http://hhi.harvard.edu/publications/applying-humanitarian-principles-current-uses-information-communication-technologies.
[15] See Nathaniel Raymond, Brittany Card, and Ziad al Achkar, "What is 'Humanitarian Communication?'" *European Interagency Security Forum (EISF),* October 2015 here: http://hhi.harvard.edu/publications/what-humanitarian-communication-towards-standard-definitions-and-protections.
[16] Ushahidi, for example, was born out of the instability of Kenya's 2008 national elections.[16] Short-code 4636 and the Standby Task Force emerged out of the need to capture and relay location-based

privacy when lives are at stake. Or to take the time necessary to think hard about the security implications of releasing real-time data to the public when vulnerable communities can and should be warned.[17]

Another challenge revolves around capacity and accountability in terms of the use of data in humanitarian operations. According to Sandvik et. al (2014), non-traditional data collection teams may be "less equipped than traditional humanitarian actors to deal with the ethical, privacy, and security issues surrounding their activities," and may not be familiar with or care about neutrality, impartiality, and accountability (key core humanitarian principles).[18]

An informed approach toward leveraging mobile technologies and remote management of local informant networks is needed. This would start with an evaluation of the 4Mi pilot to include an articulation of key thematic contributions of the pilot, an assessment of validity, reliability, ethical and security-related challenges that such an initiative faces, and a series of recommendation to improve measurement and analytical instruments and processes. These recommendations can be used to develop operational protocols and guidance frameworks for informing humanitarian actors at large how to harness mobile-data collection technologies and a network of local informants to capture granular, real-time information on drivers, profiles, conditions, and routes used for vulnerable populations on the move in low-resource, high-risk settings.

information about vulnerable populations in the wake of the 2010 Haiti earthquake. The Satellite Sentinel Project materialized out of concern over the January 2011 referendum and creation of South Sudan. The list goes on.

[17] At the Satellite Sentinel Project, we pushed forward with our mission while also advancing the methodology on how satellite imagery can be used by non-state actors for early warning and human rights monitoring in an on-going conflict setting. This meant that new methodologies (like the GRID method) were developed and it also meant that compromises on the release of data (reports) had to be made. The Satellite Sentinel Project team at the Harvard Humanitarian Initiative continues their work to advance standards and ethics as part of the Signal Program on Human Security and Technology.

[18] Sandvik et. al, 2014. Sandvik et. al (2014) also point out a range of humanitarian data privacy and security issues including the trade-off between efficiency of humanitarian action and protecting the privacy of beneficiaries in crises, the imbalance between sensitivity of data in crises contexts vs the means to protect that data, and the risks associated with humanitarian organizations becoming "unwitting" intelligence providers, as more and more data are collected from beneficiaries at greater scale, speed, and scope.

A note on legitimacy, participation, and power

When it comes to technology and big data in humanitarian action, the assumption that ICTs are unequivocally empowering or are "leveling the playing field" obfuscates matters of legitimacy and power inequities. While technology *actively generates* new roles, resource distribution, rules, relationships, and vulnerabilities in the humanitarian space,[19] there is a possibility that these effects are not inherently democratizing. Might humanitarian actors be subject to blind spots regarding populations that are overlooked by big data? Is it possible that technology itself influences the distribution of humanitarian aid in inequitable ways? Are all new actors empowered by ICTs, such as the commercial sector, "legitimate" champions of affected populations?

As humanitarians, at minimal we seek to Do No Harm: in other words, we need to make sure that our use of big data and ICTs does not either formalize current power inequities or exacerbate them. At best—and through a critical lens—we seek to harness technology and share granular real-time data in a way that empowers local and vulnerable communities to be resilient in the face of adversity, danger, or catastrophe.

Understanding matters of data governance will require the application of perspectives derived from political economy, organizational behavior, and international relations theory. As one example, the work of the Global Solutions Networks sheds light on issues related to governance and collective action through information communication technology empowered networked institutions, which requires and honest engagement with matters of bargaining power, interest and institutions, inclusion, legitimacy, representation, participation and power relationships.[20]

[19] Sandvick et. al, 2014.
[20] See Global Solutions Network: www.gsnetworks.org.

Joseph Guay is an Associate at The Policy Lab, where he leads research portfolios on “Humanitarian Innovation,” “Critical Approaches to Humanitarian Technologies & Big Data,” and “New Approaches to Humanitarian Response in Protracted Conflict and Urban Settings.” He is currently under contract with the UNHCR Innovation Lab to provide evidence-based research support for the establishment of a Global Humanitarian Lab and Global Alliance for Humanitarian Innovation (GAHI). In addition to being adjunct lecturer at Northeastern University’s Geographic Information Science Master’s program (where he developed a course on Crisis Mapping for Humanitarian Action), he is an avid humanitarian technology researcher and experienced information management capacity builder, bridging an important gap between evidence and training for impact.

Guay began his experience with humanitarian technologies at the Satellite Sentinel Project at the Harvard Humanitarian Initiative in 2011, where he provided contextual analysis for the monitoring of human rights violations in South Sudan using satellite imagery analysis. Since then, he’s been exploring the ways that information communication technologies (mobile phones, social media, remote sensing tools, unmanned aerial vehicles (UAVs)) enhance situational analysis for decision-making in early warning and response settings and how such tools fundamentally transform the very nature of the humanitarian system in the process. Recently, he is undertaking research on social stability between Syrian refugee and host communities in non-camp, urban spaces in Lebanon and Jordan, and is also currently supporting a critical research agenda on multi-modal governance in areas of limited statehood, trying to understand how transnational advocacy networks (TANs), enabled by new technologies and access, are increasingly relevant in places where official government is limited, or ceases to function according to social contract.

Guay’s has published at Brookings Institute, ALNAP, Relief Web, *The Boston Globe*, and Harvard’s Weatherhead Center for International Affairs, and he’s given talks at USGIF’s GeoInt2015 (Washington D.C.), UNHCR’s annual Partnership Consultations (Geneva, Switzerland), Oxford University’s Humanitarian Innovation Conference 2015 (Oxford, U.K.), and World Vision International’s Business for Disaster Management (B4DM) conference (Nairobi, Kenya).