The Humanitarian Data Ecosystem: the Case for Collective Responsibility

Jos Berens, The Netherlands; **Nathaniel Raymond,** United States; **Gideon Shimshon,** The Netherlands; **Stefaan Verhulst,** United States; **Lucy Bernholz**, United States

Centre for innovation, Leiden University*; Harvard Humanitarian Initiative, Signal Program on Human Security and Technology; Centre for Innovation, Leiden University; The GovLab, New York University; Centre on Philanthropy and Civil Society, Stanford University

j.b.berens@fgga.leidenuniv.nl

Abstract

The authors of this paper explore the challenges to, and potential solutions for the responsible use of digital data in the context of international humanitarian action. The goal of this paper is to present a new framework for responsible data use in this particular domain. By doing so, the paper contributes to the emerging literature on the data policy & management, introducing a data ecosystem perspective. The authors argue that 1) In digital age, the basic service provision activities of NGOs and humanitarian aid organizations have become data collection processes – they are now one and the same – and; 2) by using digital data in their operations these organisations subject themselves to a range of (domestic) laws and regulations to which they were not before. The authors conclude with the need for public policy stakeholders to invest in workable data policy practices and norms that cross organisational boundaries when designing humanitarian response mechanisms. While this is true across sectors, the characteristics of the realm of humanitarian action make developments in this space indicative of potential risks, harms and mitigation measures in other sectors.

Keywords – humanitarian action; big data; governance; responsibility; ecosystem.

Section I - Overview of Problem, Pain, and Peril: "Current approaches based on 20th century ethics and data theory"

Datafication - the phenomenon where most aspects of our (digital) lives generate digital data that can be leveraged to generate new insights into our environments and human behaviour in them - is introducing new tools, changing the way humanitarian actors go about fulfilling their mission, and is offering these actors ways to be more effective and efficient with less funds for more crises than ever before through responsible innovation.[1] At the same time, the current ungoverned and uncoordinated use by civil society of information communication technologies (ICTs) and the data they produce threatens vulnerable populations, as well as the organizations employing these tools and techniques.[2] These risks remain conflated and poorly understood, in part due to a focus on individual organisations and data sets, as separate from their surroundings and partners.

Civil society is fragmented by design. In humanitarian work, the multitude of independent organizations, government actors, multinational initiatives, individuals, private sector resource providers, and digital platforms constitute a dynamic ecosystem with no clear leader or dominant force. These different institutions, each with their own mandates, have long worked to coordinate their services and care while simultaneously maintaining institutional autonomy.

As humanitarian organizations now seek to collect, use, store, and share digital data to improve their collective results. Their institutional walls - although they might exist for good reasons to separate operations of a differing nature or with a different beneficiary cohort - are often at odds with best practices in data governance. Where it once served to have individual organizations set its norms and practices for governing human and financial resources,

---

[1] See for example, UN Independent Expert Advisory Group on the Data Revolution 2014.

[2] The Centre for Internet and Society (McDonald, S. M.) 2016.

digital data flow between and across organizations - highlighting at each crossing a new set of interdependencies. While all of civil society need to define new norms and practices for the safe, ethical, and effective use of digital data, the humanitarian sector is already well into attempting cross-institutional work while bound to institutional governance mechanisms.

Data is used for a diverse set of purposes in the humanitarian field that can range from improving situational awareness of assistance providers; connecting affected populations to response activities; and monitoring and evaluating the delivery of aid to populations. Given the need for aid organizations to coordinate both within and across specific technical subfields, this data is almost always being collected and used within a humanitarian ecosystem regardless of its area of application.[3]

Another common characteristic of humanitarian data is that it is very often being collected from, as well as applied to, highly vulnerable populations often located in intrinsically non-permissive and highly dangerous settings. Despite data being used for common purposes with similarly vulnerable populations across organizations and technical applications, there are currently no common minimal technical standards or professional ethics for humanitarian action's use of such data.[4]

The current approach to ethical and technical standards by the humanitarian data ecosystem, to the degree one exists, can best be described as a patchwork of individual institutional codes of conduct and checklists primarily addressing data use and storage.[5] These codes and standards, when they have been established within organizations, are largely based on a general and vaguely defined ethic of "doing no harm" focused primarily on the protection of personally identifiable information, or "PII", captured by a particular organization.

Thus the authors see "experimental" uses of digital data and ungoverned data sharing "pilot programs" with mounting evidence of potential legal violations and population-specific harm.[6] The challenge is deeper than simply an impulse to innovate. We are increasingly aware of the disconnect between 20th century ethical models built for individual institutions and the cross-institutional nature of digital data systems.

The existing ethical frames in the humanitarian space also presume a closed loop approach to protecting PII, when today's triangulation of different data sets requires attention to privacy protection in an ongoing, open-ended way. In short, these codes aim to "do no harm" without having common agreement on what constitutes the "harm" they aim to mitigate or prevent.[7] Thus, this "patchwork approach" has become fundamentally inappropriate and unsustainable given the interconnected nature of humanitarian data streams, as well as the shared liabilities to humanitarian organizations and affected populations inherent within these streams.

What's more, the PII-focus of the current ethical frameworks available to humanitarians does not in any way address the emerging and increasingly common type of data that humanitarians are collecting, analyzing and deploying in both natural and manmade disasters: Demographically Identifiable Information, or "DII". Raymond defines DII as follows:

> Demographically Identifiable Information, or DII, is defined as either individual and/or aggregated data points that allow inferences to be drawn that enable the classification, identification, and/or tracking of both named and/or unnamed individuals, groups of individuals, and/or multiple groups of individuals according to ethnicity, economic class, religion, gender, age, health condition, location, occupation, and/or other demographically defining factors.

> DII can include, though is not limited to, personal identifiable information (PII), online data, geographic and geospatial data, environmental data, survey data, census data, and/or any other data set that can - either in isolation or in combination - enable the classification, identification, and/or tracking of a specific demographic categorization constructed by those collecting, aggregating, and/or cross-corroborating the data.[8]

Thus, individual PII-focused, informed consent mechanisms are insufficient, and potentially harmful, when organizations are tracking time and place-specific movement/status data of large demographically delineated groups. A new lens is needed to understand how we can meet the need for 'responsible innovation' in the humanitarian sector.

Section II - Defining the Humanitarian Data Ecosystem and Understanding its Importance

---

[3] UN Office for the Coordination of Humanitarian Affairs 2016.

[4] See for example the Humanitarian Practitioners Network 2015.

[5] See Raymond, N. A. 2016.

[6] Sean McDonald, S. M. 2016.

[7] *Ad ibidem* and see Raymond, N. A. 2016.

[8] Raymond, N. A. 2016.

The implicit acceptance of the existing governance disparity has complicated the creation of an alternative governance and management framework for addressing these ethical, sustainability and operational challenges, because of the diverse nature of the actors and lack of common doctrine, objectives and ethical requirements. This paradigm has endangered the relevance and continuance of humanitarian principles.[9] We propose the fundamentals of an alternative perspective on, and approach to humanitarian data responsibility, rooted not in 20th century governance concepts, but in a systems approach to what we call the 'humanitarian data ecosystem'. Section II outlines the implications this new perspective should have for a new design of data governance.

II. A   Actors and Building Blocks

The humanitarian data ecosystem is made up of an inherently heterogeneous mix of traditional actors (NGOs, UNAs, Govs) and non-traditional actors (voluntary technical organizations, private sector, academia). Each actor has a particular set of interests, motives, assets, skills and capacities, etc. Members of the same sub-group of actors will often have a similar approach to humanitarian data use, driven by the same motives and working towards the same goals. These similarities often give rise to inter-organizational competition for funding, a dynamic that incentivises (irresponsible) experimentation with humanitarian data.

Between various sub-groups tensions are different: there are major strategic and operational variations between the types of actors using digital data. Where private sector are already using 'big data' technologies on a large scale to boost their profit, including in crisis areas, resource-constrained NGO's are often still struggling to tap the potential of 'big data' to streamline their processes and enhance their impact.

An increasing majority of NGOs now use digital systems to provide their basic services have now folded data creation/collection into their core businesses.[10] Thus, even if they don't have capacity to make optimal use of the data they collect, they face the challenges of mitigating its risks.[11]

---

[9] Raymond, N. A. and Card, B 2015.

[10] For various examples, see The New York University GovLab 2016.

[11] This is comparable to the situation in which humanitarian actors find themselves when they might not be expected to implement protection programmes, but still have the responsibility to incorporate protection principles into their programmes. See the *2012 World Vision Mainstreaming Protection Standards.*

II.B   Connections and Flows

These actors do not operate in a vacuum. Rather, they constantly engage in dynamic, ad hoc collaborations to enable combinations of skills, tools, data and platform assets, expertise and access necessary to apply ICTs and collect digital data. These collaborations generally emerge out of necessity in a given crisis, and can then be maintained. The Humanitarian Data Exchange (HDX), set up by the United Nations Office for the Coordination of Humanitarian Affairs (UN OCHA), is an example of an effort that aims to bundle connections between a variety of humanitarian actors in which they exchange data (flows).[12] Where this digital data, prior to the existence of HDX, would have been shared between humanitarian actors under bilateral agreements, these 'data flows' are now sustainably accessible to more parties.

Through these connections, flows emerge, often in regular patterns, by which assets (data and data analytics services) move through the ecosystem. By collaborating, either formally or informally, actors share value and can improve the overall efficiency and impact of the ecosystem, for example by preventing unnecessary duplication and bundling resources. However, in these flows, flaws can be also spread through the ecosystem easily. Collaboration and interdependency require a level of trust that must be adequately informed.

Together, these ad hoc partnerships create a dynamic political and philanthropic economy undergirding the humanitarian data ecosystem that is encouraged by funders, private sector agenda, and resource low/deprived state of NGOs dealing with unprecedented access/information/and delivery challenges in increasingly non-permissive, highly volatile, large geographic scale environments.

II.C   Forces and Challenges

The implicit and unavoidable acceptance of this pattern has complicated the creation of a governance and management framework for addressing these ethical, sustainability and operational challenges because of diverse nature of the actors and lack of common doctrine/objectives/ethical requirements. Without a governance framework that facilitates responsible and sustainable sharing of data, especially when data is needed urgently in crisis situations, it is not surprising that humanitarian actors seek ad hoc and easy ways to inform their partners when needed. However, as crises follow one another, these ad hoc collaborations, unsupported by any research or evaluation, form a paradigm that, in it's chaos and pragmatism, is

---

[12] *Humanitarian Data Exchange.*

endangering the relevance and continuance of humanitarian principles in digital data sharing.[13]

These ad hoc and polyglot humanitarian data ecosystems that far supersede traditional ethical frameworks and trade in experimental applications of DII collection, fusion, and deployment have two primary effects: 1) Causing all actors to be affected by risks generated by data applications of all other actors in their ecosystem; and 2) causing the overall level of risk management and ethical compliance to be determined by the mean aggregate level of data responsibility evinced by all actors in a specific operational context.

II.D   Implications

Given this paradigm, continued application of individual institutional codes of conduct and policies will inevitably fail, no matter how comprehensive they are. The interconnected nature of the risk and the collective duty for data responsibility can only be addressed through ecosystem wide ethics and governance. Even though we now diagnose these problems well, we do not yet have a clear methodology for developing such governance and for designing the mechanisms that can adapt to change when needed.

The lack of support for common data responsibility standards given by the current humanitarian civil society economy and individual institutional agendas continuing to misrecognize the development of common data governance, professional standards, and ethics, will potentially lead to a "Goma level moment" ala the failure of siloed ground response in 1994 post-Rwanda humanitarian ops but with data.[14]

Section III: Past Precedents for Ecosystem Wide Ethics in Humanitarian Action

The growing reliance by humanitarian actors on either the derivation, creation, or fusion of DII data from combinations of often non-consented data sources and platforms is generating new and different potential threats and harms to both affected populations and practitioners.Ironically, the dangers and inefficiencies of the humanitarian community's continued application of a balkanized approach to governing how it uses data and ICTs during response should be familiar to practitioners.[15]

The humanitarian response in the Great Lakes Region of Africa to the refugee crisis caused by 1994 Rwanda Genocide was a seismic event in the history of humanitarian action and resulted in major systemic reforms to the field. Because of redundant, poorly coordinated operations by multiple organizations that did not share common ethical and technical standards, the aftermath of the Rwanda response forced the humanitarian community to begin to measure itself by adoption of minimum standards across the ecosystem.

Borton summarizes the findings of the landmark Joint Evaluation of Emergency Assistance to Rwanda (JEEAR) launched to study the as follows:

> Whilst the major failings clearly lay within the political/military domains, shortcomings in the humanitarian response were also apparent. Study 3 was particularly critical of: coordination arrangements within the UN; the lack of preparation by key agencies ahead of the massive refugee influx into Goma in July 1994; and the poor performance of some international NGOs. Whilst many international NGOs had performed impressively, the unprofessional and irresponsible behaviour of some had, most probably, caused unnecessary loss of life.[16]

The JEEAR has now become recognized as a pivot point in the development of current humanitarian architecture. The impact of the study's findings on the humanitarian community led to the current frameworks and mechanisms by which the quality of response activities by both individual organizations and the broader humanitarian ecosystem are now evaluated. Chief amongst these reforms was the development of the Rights-Based Approach (RBA) enshrined in the Humanitarian Charter and the creation of minimum technical standards, which are known as Sphere Standards.[17]

The net result of the post-JEEAR reforms is the wide understanding and acceptance in the humanitarian community that the quality of a response to a disaster is now measured on the whole ecosystem's compliance with these ethical and technical standards. A similar normative, ecosystem-wide framework based on measuring the quality of overall data governance process is now required for the humanitarian ecosystem's use of digital data and ICTs.

The process by which this normative framework for the humanitarian ecosystem's use of digital data and ICTs is established must address several specific issues unique to the humanitarian context, including the following:

---

[13] Raymond, N. A. and Card, B 2015.

[14] Raymond, Howarth, Hutson 2015.

[15] Such as during the West-African Ebola Outbreak in 2014-2015, as mentioned above in Section I. See The Centre for Internet and Society (McDonald, S. M.) 2016.

[16] Humanitarian Practice Network (Borton, J.) 2014.

[17] The Sphere Projet, *Humanitarian Charter and Minimum Standards in Humanitarian Response.*

- Core humanitarian principles, including humanity, neutrality, independence and impartiality, must be translated and applied within the context of current data uses as part of this framework;[18]
- The rights of individuals and populations in the context of the use of their data in humanitarian action must be identified and enshrined;[19] and
- Minimum technical and ethical standards for what constitutes the professional and appropriate methods and settings for the use of specific tools and approaches must be identified and agreed.[20]

Section IV: Achieving Accepted Governance Structures for the Humanitarian Data Ecosystem

There are significant obstacles to accomplishing these critical prerequisites to an ecosystem wide framework for data governance in current humanitarian action. Some of these obstacles include a lack of agreement on who is included in the humanitarian data ecosystem, particularly private sector actors and voluntary technical organizations; an absence of funding for this work; and a low appetite for additional forms of self-regulation among some sectors of the humanitarian field. However, the humanitarian data ecosystem must address these issues if it hopes to avoid repeating the mistakes of its past.

We need to develop governance processes for the humanitarian use of digital data that are simultaneously ecosystem focused, collectively created and used, and which can be applied by all vested parties throughout the lifecycle of data. This will require a recursive approach - to developing the ethical practices and the methods of applying them simultaneously. We can assume that digital data and infrastructure will continue to play a foundational role in the future of humanitarian action. The viability of the humanitarian sector itself depends on its ability to govern and manage digital data as both assets and liabilities.

All parts of the humanitarian data ecosystem - from funders to beneficiaries - need to participate in creating the principles and practices for the ethical use of digital data. We can learn from the small but growing body of research, consider the role of incentives, and borrow from adjacent sectors (e.g. medical research, human rights) where work is underway to define community consent, obligations to use,

and experimental guidelines. An initial agenda for action would include delineating the common lifecycle of digital data use in humanitarian work, scaling data interventions according to existing evidentiary bases, and defining the core actors within the ecosystem and identifying ways they can engage in the process. Importantly, the guiding frame for humanitarian data use should be clear both on rights for affected populations, as well as the corresponding obligations for humanitarian actors looking to use their data.

Conclusion

The authors first outlined the current state of affairs in humanitarian use of digital data: data is collected, stored, and used by an uncoordinated, diverse group of actors with varying degrees of expertise in responsibly leveraging digital data for humanitarian purposes. The lack of alignment causes a range of inefficiencies, but also risks that can spread easily across organisations.

In section II, the authors proposed an alternative view on this use of digital data: that of complex data ecosystem comprised of a variety of actors that are touched by flows of digital data due to data sharing and data-related service delivery, and hence become part of a phenomenon that stretches beyond their organisational boundaries. Further, there are forces and challenges on the same meta-level – including but not limited to competition for funding, resource constraints throughout the humanitarian field, and political pressure to, and attention for organisations that do, innovate faster.

These circumstances lead to several implications, both for individual humanitarian organisations and for the humanitarian data ecosystem as a whole. For individual organisations, the most important consequences of their participation in this ecosystem are to 1): look beyond the role they take up in the 'data-lifecycle' and consider previous and following steps and roles, 2): develop sound data responsibility strategies not only to prevent harm to their own operations but also to other organisations in the 'data-lifecycle', and 3): to collaborate with and learn from other organisations, both in the humanitarian field and beyond, to establish broadly supported guidelines and standards for humanitarian data use.

For the humanitarian data-ecosystem as a whole, and for coordinating organisations, the implication is that their monitoring role should shift from a focus on individual organisations, to data life-cycles. Streams of data will need to be tracked and the way they are leveraged should be better understood based on the ecosystem view and specifically on the forces and challenges outlined in section II.

---

[18] Raymond N. A. and Card, B. 2016..

[19] International Peace Institute Global Observatory (Raymond N. A.), 2013.

[20] Humanitarian Practitioners Network (Raymond N.A. and Harrity, C. S.) 2016.

Such a more holistic framework for data policy can inspire organisations in the sector to strengthen their outlook on responsible data use, by going beyond their internal data life cycles. As the data revolution unravels, humanitarian actors should therefore invest in cross-organisational understanding of data use, with the aim to avoid inappropriate - and potentially harmful - (re-)use of data sources.

References

UN Independent Expert Advisory Group on the Data Revolution 2014. *A World That Counts*, available from: http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf.

The Centre for Internet and Society (McDonald, S. M.) 2016. *Ebola - A Big Data Disaster*, available from: http://cis-india.org/papers/ebola-a-big-data-disaster.

UN Office for the Coordination of Humanitarian Affairs 2016, *Building Data Responsibility Into Humanitarian Action*, available from: https://docs.unocha.org/sites/dms/Documents/TB18_Data%20Responsibility_Online.pdf.

Humanitarian Practitioners Network 2015. *Addressing the Doctrine Gap: Professionalising the Use of Information Communication Technologies in Humanitarian Action*, available from: http://odihpn.org/magazine/addressing-the-doctrine-gap-professionalising-the-use-of-information-communication-technologies-in-humanitarian-action/.

Raymond, N. A. 2016. Beyond "Do No Harm" and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society's Use of Data. In L. Taylor, L. Floridi, & B. van der Sloot (Eds.), *Group Privacy: new challenges of data technologies*. Springer.

McDonald, S. M. 2016. "Ebola: A Big Data Disaster," The Centre for Internet and Society, March 2016, available from: http://cis-india.org/papers/ebola-a-big-data-disaster.

Raymond, N. A. and Card, B 2015. *ICTs and Humanitarian Principles White Paper*, available from: http://hhi.harvard.edu/sites/default/files/publications/signal_program_humanitarian_principles_white_paper.pdf.

The New York University GovLab 2016. *Selected Readings on Big Data and Humanitarian Response*, available from: http://thegovlab.org/data-and-humanitarian-response/.

2012 World Vision Mainstreaming Protection Standards, available from: http://reliefweb.int/sites/reliefweb.int/files/resources/Full_Report_3752.pdf.

Humanitarian Data Exchange, available from: https://data.humdata.org.

Raymond, Howarth, Hutson, 2015. *Crisis Mapping Needs an Ethical Compass*, available from: http://globalbrief.ca/blog/2012/02/06/crisis-mapping-needs-an-ethical-compass/.

Humanitarian Practice Network (Borton, J.) 2014. *Twenty Years On: The Rwandan Genocide and the Evaluation of the Humanitarian Response*, April 2014, available from: http://odihpn.org/blog/twenty-years-on-the-rwandan-genocide-and-the-evaluation-of-the-humanitarian-response/.

The Sphere Projet, *Humanitarian Charter and Minimum Standards in Humanitarian Response*, available from: http://www.spherehandbook.org/en/the-humanitarian-charter/.

Raymond N. A. and Card, B. 2016. *Applying Humanitarian Principles to Current Uses of Information Communication Technologies: Gaps in Doctrine and Challenges to Practice*, available from: http://hhi.harvard.edu/publications/applying-humanitarian-principles-current-uses-information-communication-technologies.

International Peace Institute Global Observatory (Raymond N. A.), 2013. *Is There a Human Right to Information During Disasters?*, available from: https://theglobalobservatory.org/2013/08/a-human-right-to-information-during-disaster/.

Humanitarian Practitioners Network (Raymond N.A. and Harrity, C. S.) 2016. *Addressing the Doctrine Gap: Professionalising the Use of Information Communication Technologies in Humanitarian Action*, available from: http://odihpn.org/magazine/addressing-the-doctrine-gap-professionalising-the-use-of-information-communication-technologies-in-humanitarian-action/.