Introduction

Modern technology and innovations constantly transform the world. This also applies to humanitarian action and development aid, for example: humanitarian drones, crowd sourcing of information, or the utility of Big Data in crisis analytics and humanitarian intelligence. The acceleration of modernization in these adjacent fields can in part be attributed to new partnerships between aid agencies and new private stakeholders that increasingly become active, such as individual crisis mappers, mobile telecommunication companies, or technological SMEs. These partnerships, however, must be described as simultaneously beneficial as well as problematic. Many private actors do not subscribe to the humanitarian principles (humanity, impartiality, independence, and neutrality), which govern UN and NGO operations, or are not even aware of them. Their interests are not solely humanitarian, but may include entrepreneurial agendas. The unregulated use of data in humanitarian intelligence has already caused negative consequences such as the exposure of sensitive data about aid agencies and of victims of disasters. This chapter investigates the emergent governance trends around data innovation in the humanitarian and development field. It takes a look at the ways in which the field tries to regulate itself and the utility of the humanitarian principles for Big Data analytics and data-driven innovation. It will argue that it is crucially necessary to formulate principles for data governance in the humanitarian context in order to ensure the safeguarding of beneficiaries that are particularly vulnerable. In order to do that, the chapter proposes to reinterpret the humanitarian principles to accommodate the new reality of datafication of different aspects of society.

The impact of modern technology and digital innovations transform the way international relations in general and humanitarian and development aid in particular are conducted. For example, humanitarian drones and private satellites collect data that previously was a prerogative of governments. These developments have come so far that even academic organisations can attain satellite data from private enterprises granular enough to uncover war crimes and crimes against humanity, like demonstrated by Harvard Humanitarian Initiative’s Satellite Sentinel Project. Crowd-sourcing contributes to another
layer of crisis mapping providing local information unavailable before. This development started to take off around 2010 and was used to inform humanitarian planning just as much as election monitoring. Additionally, telecommunication companies contribute their wealth of data to development challenges and provide what they call Data for Development (D4D) or data donorship.

The acceleration of modernization in the aid sector can in part be attributed to new partnerships between aid actors and new private stakeholders that increasingly become active in the humanitarian and development field, such as individual crisis mappers, telecommunication companies, or technological multi-national corporations and SMEs. This also leads to a trend that much of the available applied research does not come from the academic community but is produced by international organizations, NGOs, think tanks and tech companies.1 This development has given many additional tools for aid operations helping organisations to improve logistics, make processes more efficient and better manage resources in humanitarian disasters.

These new partnerships between aid actors and the commercial field, however, must be described as simultaneously beneficial as well as very problematic. Many private actors do not subscribe to principles guiding humanitarian action (humanity, impartiality, independence, and neutrality). Furthermore, data is inherently political and its sensitivity changes from context to context. For-profit organizations might pursue an entrepreneurial agenda, which is problematic considering that their prospective market consists of the most vulnerable people of the world and whose vulnerability is the corporations’ entry ticket into a new market of the future. In other words, what starts as charity soon could become hard and cold business.

This chapter provides an overview over the political dilemmas that emerge from the digitalization trend in the humanitarian and development field. Section one outlines the changes in these fields initiated and driven by an increasing reliance on digital data. The next section (section two) delves deeper into the questions surrounding the political nature of data, including data vulnerability and (group-) privacy. Section three investigates the political consequences of datafication in the humanitarian and development aid and suggests the application of principles of humanitarianism also to data-driven operations and innovation as a mode of normative governance.

---
1 While this chapter aims to give an illustrative overview, given the multitude of current initiatives and the speed of development in this field, it is by necessity outdated upon publication and can only hope to remain of relevance concerning the principal issues and themes it raises.
Data Innovation in Humanitarian Action and Development Aid

Humanity means to “prevent and alleviate human suffering wherever it may be found. Its purpose is to protect life and health and to ensure respect for the human being. It promotes mutual understanding, friendship, co-operation and lasting peace amongst all peoples.”

This is the guiding principle of international humanitarianism, its own categorical imperative. From there onwards things get more complicated. On the one hand, many humanitarian organizations try to be as apolitical as possible. On the other hand, crises are seldom apolitical themselves; and in order to solve them, development aid needs to be inherently political. This means, if for example a humanitarian crisis is not solely the result of a natural disaster but the result of a combination of corruption, resource mismanagement, and a lack of disaster risk reduction planning (DRR), then development aid will have to implement good government practices, reduce corruption and implement measure that prevent natural hazards to ever such a disastrous scale requires the help of the international community. In other words, if the international community tries to get at the root causes of conflicts and natural disasters, humanitarian aid is not enough, but long-term solutions through development aid are necessary. This is when things become political.

The same goes for data. From the outset, one might think of it as neutral and passive, but already the means of its collection introduce structural biases and the utility of datasets introduces another political layer. With structural biases we meant that data, particularly when sampled representatively, will lay bare structural biases that a society has vis-à-vis one or more of their social groups. How such information in turn is used reveals even more about the political biases of a society. At the same time, data is incredibly important when it comes to the planning and implementation of crisis response and development aid. The hope for faster and more effective response mechanisms in an austere climate has accelerated the use of digital technologies over the past decade. Resources must be made use of as efficiently as possible and project implementation, for which aid organisations have to justify themselves vis-à-vis donors, has to not only satisfy economic controlling but also political agendas such as gender equality and non-discrimination. Digitalisation and computer aided logistics are only a small part of the promise of the digital era which aims at making processes more efficient. For example, the 2010 earthquake left Haiti devastated

---

and a huge number of people died or became homeless. However, this time aid was delivered predominantly aided by digital technologies such as crisis mapping through services such as Ushahidi. This lead to the emergence of the concept of digital humanitarianism.\(^3\) It marked a shift in how digital technologies would be employed in the future, such as read-write web, big data analytics, participatory mapping, crowdsourced translation, social media, and mobile technology.\(^4\) It allowed for the first time to better digitally map where aid was needed, which organisations were in the vicinity, what resources could be allocated and how to best reach people in need.

Thus, data is not restricted to the industrialized West. Crises result in huge volumes of data, like text messages from victims to their relatives, movement data of mobile phones between the cell towers, local and regional data from social media sources, and publicly accessible satellite data. Citizens, journalists, aid organizations, governments, international organizations, produce such large quantities of “big crisis data”, sifting through what is needed and reliable presents a challenge in itself. In the case of Haiti, this problem was approached by using techniques such as crowd sourcing in order to provide humanitarian operations with nearly real time crisis maps.\(^5\) One of the first such services used was called Ushahidi. This online platform allows users to upload geolocated information with mobile phones to the servers which translate the information in near real-time to maps. Subsequently, the Ushahidi technology was used to improve the election monitoring in Kenya 2013 (Uchaguzi project).\(^6\) Furthermore, map overlay technology empowered other crisis mappers (e.g. Sahanna, Standby Task Force etc.) and allowed private citizen to contribute as first responders to crises. The umbrella organization, Digital Humanitarian Network (DHN) for instance, gathers organizations to provide technological, statistical and data science expertise to humanitarian organizations and develops user-friendly tools for volunteers.\(^7\) Also, the United Nations Office for the Coordination of Humanitarian Affairs (UN-OCHA) responded with establishing the Humanitarian Data Exchange (HDX) providing already more than 5000 datasets and analyses on all sectors of humanitarian

---


5 Ibid.


action.\textsuperscript{8}

The crisis in Haiti transformed the way crisis response and intelligence needs were thought of and led to a sort of distributed humanitarian intelligence referred to as \textit{Disaster Response 2.0}.\textsuperscript{9} The 2013 World Disasters Report dedicated to the topic “humanitarian information and communication technologies (HCIT)” indicates the tremendous importance that this topic has gained in a few years after the Haiti earthquake.\textsuperscript{10} This development has only accelerated in the recent years and has encouraged the United Nations to create the Centre for Humanitarian Data and its Humanitarian Data Exchange (HDX). It has further led the World Food Programme to fund its own innovation accelerator and many non-governmental organisations, such as World Vision International, to increasingly use digital technologies from Data Analytics to Blockchain technology.

Big Data and Artificial Intelligence have become buzzwords in innovation. The reason can be found in the increasing amounts of data that are being produced daily, which call for new ways of analysis.\textsuperscript{11} Much of the data comes from the proliferation of smart phones and sensors, the Internet of Things (IoT) and other modern technologies such as wearable devices. These provide new tools for the collection of data about human behaviour, pattern recognition and forecasting. The daily petabytes produced by the digital trail that people leave behind on social media and the internet at large is just as informative as the call details record (CDR) of mobile phone users. The term “Big Data” commonly refers to massive sets of unstructured data (as opposed to structured data sets which are purposely collected), which cannot be stored in organized, relational databases.\textsuperscript{12} The multitude of sources and formats that feed into this data pool (videos, text messages, sensor data, etc.)

\textsuperscript{8} OCHA, “Ebola Crisis Page - Humanitarian Data Exchange,” accessed March 30, 2015, https://data.hdx.rwlabs.org/ebola: This includes for example: databases aggregating the number of Ebola cases and infected aid workers, water sources and opinions about water quality in Kenya, to a database of total uniformed personnel contributions of each contributing country by month, by type (troop, police, or expert/observer) and by mission.


allows gaining new insights through correlational analysis, pattern recognition and other forms of machine learning. This development has pointed people to even proclaim the end of causality and its substitution by correlational evidence. At the same time, critique of big data research devoid of theory in the social sciences and in international relations has been voiced. Since the dawn of the scientific method, first theorized by Aristotle in his Metaphysics, the role and importance of science has been to elevate our understanding from the specificities of the concrete situation and the expertise of the practitioner to the generalisation if insights into the larger mechanisms at play. In other words, the essence of science is theory building and testing. In return, science enables us to use these theories on a variety of cases to solve problems that, while being unique, are similar enough to exhibit the same properties as problems in other contexts. In short, we don’t have to reinvent the wheel over and over again. This critique of the end of causation in favour of a correlational approach is specifically targeted at the problem that without understanding causal relationship meaningful interventions and prevention efforts cannot be developed. For humanitarian action, development aid, conflict resolution and peace building to follow the learning by correlation strategy would mean that we would have to experiment every time anew on vulnerable people until we would find the solution that works.

Recently, the term “Big Crisis Data” emerged to signify big data collected during crisis and by crisis responders. Just like Big Data, we can differentiate between structured and unstructured data, whereas most of it is usually unstructured data. Given the urgency that a sudden onset crisis present, one of the biggest challenges remains the swift extraction of structured information from unstructured sources such as situation reports, mission briefings and information deriving from social media and news outlets. In general, one can structure Big Crisis Data into several enabling technologies, which call for different data sources and allow for specific data driven solutions in humanitarian crises (see table 1).

<table>
<thead>
<tr>
<th>Enabling technologies</th>
<th>Example(s)</th>
<th>Example scope(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data source:</strong> data exhaust</td>
<td>Call detail records (CDRs)</td>
<td>Epidemiology (disease spread), logistics</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Data source: online activity</th>
<th>User-generated data (e.g., emails, text messaging and social media activity)</th>
<th>Sentiment analysis, opinion mining, search and rescue</th>
</tr>
</thead>
<tbody>
<tr>
<td>The internet, machine learning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data source: remote sensing technologies</th>
<th>Satellites, UAVs, sensor networks (including mobile phones), wearable devices</th>
<th>Social behaviour and environmental analysis (e.g. earthquake early warning)¹⁵</th>
</tr>
</thead>
<tbody>
<tr>
<td>IoT, data visualization, machine learning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data source: sensing technologies</th>
<th>Wearable devices, social media</th>
<th>Healthcare (e.g., personalized medicine)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IoT, the internet, machine learning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data source: public data</th>
<th>Governmental data (e.g., census, public health and transportation data); IGO and INGO data, HDX data; monitoring and evaluation data</th>
<th>Evidence based policy making; decision making dashboards for early warning / early action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open source and open data, machine learning, blockchain technology</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data source: crowdsourced data</th>
<th>Crowdsourcing platforms (e.g., Ushahidi)</th>
<th>Crisis maps, GIS analysis and services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowdsourcing, Data visualization, neo-geography, machine learning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Big crisis data sources and enabling technologies¹⁶

The Humanitarian Open-StreetMap project also gained importance during the Haiti earthquake, as it became the sources of mapping data for most of the UN agencies responding


to the crisis. Furthermore, a related project is the Missing Maps, which focuses on mapping the most vulnerable places and regions in the developing world. Given the fact it makes openly available highly detailed maps of previously unmapped regions, it has become one of the go-to sources for humanitarian organizations enabling them to improve their logistical planning and supply chain management. Crisis mapping has ushered in a new interdisciplinary field of neocartography that increasingly also drives research. This trend has further been enhanced by the availability of up-to-date satellite imagery by companies such as DigitalGlobe. Crowdsourcing and crowd computing techniques support the analysis and annotation of these maps.

The utility of geographic information systems (GIS) and neocartography is even used in peace building efforts and international criminal justice. Already in 2011, the Satellite Sentinel Programme of Harvard Humanitarian Initiative (HII) used location data correlated with surveys, photos, maps and interviews and helped finding eight mass graves in Sudan. Ever since, the expertise and service provided by HII has aided many UN and non-governmental agencies in the fields of peace and justice as well as humanitarian action and development aid. This and similar advancements in crisis analytics raise hopes that we will be able to predict conflicts more accurately and long-term and maybe even stop wars before they happen.

From the perspective of Data for Development (D4D) an important facilitating principle is the move towards open government data to increase transparency and to allow private access and research in the field of development and good governance previously not possible. Furthermore, the political move towards open research was aimed at accelerating soft-power transfer to developing nations to maximize the social and economic benefits of research. To that end, in 2013, the Group of 8 (G8) Science Ministers issued a joint

---

17 See: https://www.hotosm.org
18 See: http://www.missingmaps.org
statement promoting open research and open research data.\textsuperscript{23} As suggested by Ali et al., “Open data also promotes a culture of creativity and public wellbeing as is evident by different hackathons that are being organized to tap the potential of open data in terms of useful mobile applications (e.g., the local government of Rio de Janeiro has created the Rio Operation Center aimed at harnessing the power of technology and big data to run the city effectively in terms of transport management, natural disaster relief, mass movement and management of slum areas).”\textsuperscript{24}

Also in the field of development, the United Nations joined the trend of datafication and data driven innovation and set up the UN Global Pulse in 2009, with the purpose of harnessing data for human development.\textsuperscript{25} Associated with UN Global Pulse is a network of Pulse Labs all over the world bringing together people from different fields and professions to support innovation with data for humanitarian and development affairs. In order to do that effectively, these labs rely on “data philanthropy”, i.e. donation of data sets, for instance call detail record by mobile phone companies. As an illustration of its research, the Pulse Lab Kampala developed together with academic and non-academic partners a tool for data visualisation and interactive mapping to support actors in the field of development and disaster response to predict disease outbreak on the basis of movement patterns generated by analysing how people moved between cell towers.\textsuperscript{26}

One important point of contention with the concept of “data philanthropy” or data donorship needs to be mentioned. Much of the data that companies so generously donate has been collected from the users that are affected by the crisis. In other words, were it not for the users who produce the data and who have to transfer the rights to their personal data to these data companies (i.e. telecom, social media, banking institutions etc.) these companies would have nothing to donate. This begs the question whether it would not be fair to impose an obligation to data sharing of crisis relevant data collected from crisis affected populations on these corporate agencies. Of course, the data recipient would have to be a data broker beyond reproach, who could be trusted and controlled never to misuse

\textsuperscript{26} See: Pulse Lab Kampala, Response to Disease Outbreak, http://diseaseoutbreaks.unglobalpulse.net/uganda/
such data and to release such data on a need to know basis to aid agencies in the event of emergencies. The United Nations Office for the Coordination of Humanitarian Affairs and its Centre for Humanitarian Data could for example take on such a role.

The Political Nature of Data

While the use of Big Data in humanitarian and development aid promises huge improvements in efficiency of interventions, it also has the potential for misuse. As such the nature of data is not neutral. On the one hand, data is a dual use object that, depending on the context, can either help or put people at risk. For example, in a political crisis, the location data of actors can be used for the purpose of military targeting putting people at risk. The very same location data can be crucial to find victims of earthquakes and other natural disasters. Crisis mappers had to learn this lesson already in 2010, when the project Pakreport.org tried to support crisis response to a food and flood crisis in Pakistan. After the project had crowdsourced digital maps and provided them open access, Pakistan based Taliban forces threatened to attack World Food Programme and Médecins Sans Frontières teams operating in their region. As a consequence, crisis mappers and other humanitarian innovators are nowadays increasingly aware of the potential danger and develop data sharing policies in order to ensure that sensitive data is not being shared.

On the other hand, the way in which Big Data and personalized information is being handled not only changed the way marketing is being done. Even individuals’ search results are tailored to their online browsing habits. While personalized information using Big Data can be a service, it also can result in echo chambers and might even be misused for manipulation practices, such as Big Nudging. An echo chamber can for example be the result of personalized information provided by search engines that is tailored to a user’s interests. While this can be a service to help find what the user is interested in, it might also reinforce established opinions rather than offer alternative views. Regarding Big Nudging, a recent example has been the revelations about Cambridge Analytica. This campaigning

---


28 Mireille Hildebrandt, Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology (Cheltenham, UK: Edward Elgar Pub, 2015), 90.

company, which among other clients served the Brexit campaign and the Trump presidential campaign, used personal data on individual voters to create psychographic profiles. These profiles were in turn used to tailor campaign messages in order to trigger voters fears to elicit a desired voter behaviour. Beyond that, Big Data is generating insights not only about individuals, but also about groups. Specifically, in development and humanitarian contexts, this creates new types of challenges and vulnerabilities.

In 2014, the research consultancy Caerus Associates in collaboration with First Mile Geo, which provides online data collection, analysis and sharing services, mapped the dynamics of the conflict in Aleppo. At that time the industrial capital of Syria was one of the most affected cities in the Syrian conflict and the battle ground of the Syrian government forces, militia and rebel parties. Caerus Associates aimed to map the conflict zone in an effort to provide information to stakeholders, such as humanitarian organizations to allow the flow of aid into regions of need. From September 2013 to January 2014, the mapping was conducted by a team of locals equipped with GPS devices and connected to cloud storage. The data Caerus collected included geolocation, political opinions about warring factions, presence of bakeries, electricity etc. While the work of Caerus is certainly important, the consultancy released its information with highly sensitive information potentially endangering surveyed citizen already in February 2014 as open data. The actual needs assessment for humanitarian purposes the study claimed to provide was rather insufficient for operations planning and it remains doubtful to what extent the benefits outweighed the risks. In general, there is a tendency both in development and humanitarian aid that collecting more data is better. This is paired with the political idea that open data improves transparency. While this might in general be the case, in the specific contexts of political and social volatility in which many humanitarian and development actors find themselves, open data can be extremely dangerous, both to aid workers as well as to beneficiaries. This illustrates an ongoing worrying trend regarding data disclosure practices, where the idea of open data is treated rather uncritically and almost as an ideology.

Privacy concerns remain high on the agenda of data protection experts. However, with the ascent of Big Data, another dimension of privacy has to be added as a risk factor – group privacy. This form of privacy concerns aspects of vulnerability of groups, the disclosure of which might increase the group’s exposure to repression, manipulation and

other negative consequences. In general, this kind of data concerns demographics connected with geographical information. For example, Linnet Taylor details the risks that the use of CDR in a data for development challenge conducted by the mobile phone company Orange in Côte d’Ivoire introduced despite the company’s stringent rules of privacy protection.\footnote{Linnet Taylor, “No Place to Hide? The Ethics and Analytics of Tracking Mobility Using Mobile Phone Data,” Environment and Planning D: Society and Space, October 6, 2015, 0263775815608851, doi:10.1177/0263775815608851.} In general, privacy and anonymization cannot be understood as binary (either anonymous or not), but have to be seen on a continuum: on the one end of the continuum, the database contains all private and individual data in a granular fashion; on the other end, the database is completely anonymous to the degree that it contains very little to no information. The analyst thus has to make a choice between how much information to retain while at the same time safeguarding the privacy of individuals. This principle was already formulated by Dalenius in 1977 for privacy protection in statistical databases: “nothing about an individual should be learnable from the database that cannot be learned without access to the database”.\footnote{Tore Dalenius, “Towards a Methodology for Statistical Disclosure Control,” Statistik Tidsskrift 5 (1977): 429–44; Cynthia Dwork, “Differential Privacy,” in Automata, Languages and Programming, ed. Michele Bugliesi et al. (Berlin, Heidelberg: Springer, 2006), 1–12.} This becomes even more problematic with the use of large scale databases where \( n \) approaches the whole population (\( n=all \)). Such databases contain much data that can be considered sensitive from a group privacy perspective. Whereas, many humanitarian and development organizations have some policy regarding privacy protection, hardly anyone concern themselves with the shift from personally identifiable information (PII) to demographically identifiable information (DII).\footnote{Nathaniel A. Raymond, “Beyond ‘Do No Harm’ and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society’s Use of Data,” in Group Privacy (Springer, 2017), 67–82, http://link.springer.com/chapter/10.1007/978-3-319-46608-8_4.} The Netherlands Red Cross and Red Crescent movement’s tech and data start-up called 510 has recently released first of its kind data protection guidelines that include issues such as group privacy and DII protection.\footnote{See: https://www.510.global/510-data-responsibility-policy/}. 

Data Ownership also remains a problem in humanitarian crises and also in general. Currently, there are no legally binding norms that would regulate the data use of humanitarian and development aid organisations in their international conduct. To redress this problem, the Signal Code produced by Harvard Humanitarian Initiative distilled four human rights based data principles to guide aid agencies in their operations.\footnote{These principles are: 1. The Right to Information; 2. The Right to Protection; 3. The Right to Data Privacy and Security; 4. The Right to Data Agency; 5. The Right to Redress and Rectification. Faine Greenwood et}
interest to data ownership is principle 4: “Everyone has the right to agency over the collection, use, and disclosure of their personally identifiable information (PII) and aggregate data that includes their personal information, such as demographically identifiable information (DII).” This further means that populations have the right to be informed about data related activities of aid agencies, which concern their own data, during all phases of information acquisition and use.

The right to data agency is one that has much promise even beyond the humanitarian field, also in development and international human rights law. Data agency rights encompass the right to protection from non-consensual experimentation, for example. The issue of experimentation is a constant concern for humanitarian operations. How can one try out new tools of data collection, data processing and analysis as well as innovations in technology (e.g. humanitarian drones) during crisis without risking the lives of vulnerable people if the experiment fails? Here the role of collaboration between development aid and humanitarian action (i.e. LRRD – linking relief, rehabilitation, and development) cannot be overstated. Since the success of implementation of new technologies and data-driven solutions is contextually bound, experiments cannot be conducted outside of this context, i.e. mobile phone penetration, technology adaptation, internet access, cultural practices etc. strongly inform the success and failure of such innovations. Even in non-crisis situations, e.g. the Facebook mood study, the lack of informed consent and sub-standard research ethics in Big Data driven innovation can at best lead to public outrage and at worst to dramatic negative consequence for research subjects. In order to guarantee that informed consent and research subject participation as well as notification of data collection and uses can be sufficiently implemented, the humanitarian sphere needs to closely cooperate with the development field with regards to innovation. This further means, that to ensure the integrity of humanitarian agencies, the humanitarian principles, a set of norms that guide most of humanitarian and relief efforts to create an apolitical, neutral, operational space, need to


36 Ibid., 16–17.


extend to some degree also to development aid, when it comes to data driven innovation (see section 4).

The general debate on data ownership can learn much from the ideas developed by the Signal Code and the problems that are more pronounced in the domain of humanitarian action due to volatile contexts. In order to be able to develop further regulations and governance mechanisms regarding personal data and to prevent the misuse of such (think for example of Cambridge Analytica), it is necessary that we develop a clearer view on the fundamental question of data ownership; i.e. who the data belongs to and what aspects of that ownership can be under which circumstances waived or contracted out and which aspects can never be waived contractually. This will be important to curb the current practices of companies which determine via non-negotiable terms that they can do what they please with the personal data of users.

On a macro-level, Broers and Taylor identify two ways in which the practices of corporations, which can largely be labelled information capitalism, change power structures in low and middle-income countries: “First, the way in which practices of informational capitalism are shaping development interventions via public–private partnerships. Multinational technology firms are involved in interventions to map, sort and categorise, often also connecting people to new technologies for the first time, and in return capture data on new technology users and, through that data, new consumers and markets. The second trend is for the visibilities created by these corporate interventions to constitute new population-level data-bases and maps.”

They characterise such newly created data infrastructures as “shadow maps”, which double state data and which operate in parallel to state data collection infrastructures. This gives rise to distributed forms of governance and empowers non-state actors (specifically corporate data brokers) and foreign state actors to intervene in otherwise sovereign affairs.

Humanitarian Principles in Data-Driven Innovation and Operations

The humanitarian principles (humanity, impartiality, neutrality, independence) are a set of norms, which together regulated the conduct of humanitarian agencies to ensure that outside

---

40 This section is based on a presentation about "Humanitarian Intelligence and Data Protection" given at the European Parliament for the European Data Protection Supervisor on 18 May, 2017, at the European
stakeholders, including warring factions and parties to a conflict, would perceive aid organizations as neutral vis-à-vis the goals of the conflict parties and foreign political agendas. To some degree these principles are even legally binding to UN agencies operating in humanitarian crises. The humanitarian principles are, however, not binding to development agencies, which by their very nature have political objectives. This section translates the humanitarian principles to data-driven operations. At the same time, it will become clear why in many cases these principles would also increase ethical responsibility in data innovation on behalf of humanitarian innovation and data collection done in the development sector.

Humanity, the first and most important principle, is also referred to as humanitarian imperative: aid should be given on the basis of need alone. Neutrality is understood as “a duty to abstain from any act which, in a conflict situation, might be interpreted as furthering the interests of one party to the conflict or jeopardizing those of the other.” Impartiality is described by the ICRC as: “No discrimination as to nationality, race, religious beliefs, class or political opinions. It endeavors only to relieve suffering, giving priority to the most urgent cases of distress”. And independence means maintaining autonomy and being sovereign in making “decisions, acts and words”. Independence is maintained by aid organizations when they do not allow “any intrusion of politics into their own sphere of action.” This intrusion can also come from economic fields and business interests. Given the dependence of development and humanitarian actors on tech companies to support innovation efforts, what might look at the outset as philanthropism and corporate social responsibility (CSR) on behalf of corporate agents might often cover up a hidden agenda of increasing market share.

Humanity means that the humanitarian imperative comes first. Currently, the trend is to collect and retain as much data as possible in the hope future data analytics of “legacy

---

43 Pictet, “The Fundamental Principles of the Red Cross.”
45 Pictet, “The Fundamental Principles of the Red Cross.”
data” will help understand conflicts and disasters better. The specific vulnerabilities of people in need might require, however, a very restrictive data policy for their own protection. Specifically, in the field of health data such a policy should comply with medical standards of restricted data use, sharing and deletion. For humanitarian Big Data, this might mean more precisely that the collection of data about a vulnerable set of people should not be random but strictly governed by a “need to know” principle. This has to do with the nature of vulnerability in crisis situations. Vulnerability decreases data agency of people affected by disasters already by the mere fact that their priorities shift towards survival. Furthermore, in political crises, any data collected increases the vulnerability as it can be misused by warrying factions to target the beneficiary population when data breaches occur. In addition, (big) crisis data that contains PII and DII should only be shared on a need-to-know basis, since data sharing also increases the exposure risk of the data. In general, the principle of humanity in conjunction with the increased vulnerability of people also requires that humanitarian organizations comply at the very least with standards set in academic research ethics in the fields of social and life sciences.

Neutrality means that humanitarian actors must not take sides in the conflict. Concerning Big Data collection and analysis this posits the problem that data needs to be collected in such a way that it represents actual need and not for example the access to internet or cell phones which might be different between the parties of a conflict. Similarly, neutrality means that aid should be delivered without discrimination as to nationality, race, religious beliefs, class or political opinions. This puts a very strong and directional requirement on the Big Data principle of veracity. More specifically, this means that data collection, processing and analytics should remain representational and not increase the vulnerability of populations by including or perpetuating biases or worsen the exposure of vulnerable groups by emphasising differences contained in DII. This raises the important question to what extent aid agencies are required to act upon information they uncover about structural biases. For example, when a group is identified to be structurally disadvantaged by a law or policy to the degree that this would violate the human right of non-discrimination, would the aid agency be obliged to correct that bias either by implementing non-discrimination elements into the operation? This question remains quite problematic.

46 I.e. type of data, categories used, data sensitivities assessed and taken into account with regards to the vulnerability of specific groups such as women, children, minorities, culturally or politically sensitive issues.

because it would mean for humanitarian action to become political.

Finally, independence means that “humanitarian action must be autonomous from the political, economic, military or other objectives that any actor may hold with regard to areas where humanitarian action is being implemented”. This last element might possibly be a problematic stumbling block as Big Data collection in many instances will happen through corporate actors. Big Data collection of CDR data, mobile phone cash transfer data is conducted by telecommunication service providers and financial institutions with commercial interest. The result of the data analytics might expose vulnerable groups as target groups, or micro markets, for specific tailor-made services from which the service provider might want to profit. At the same time, corporate agencies can provide crucial data as “charity” to humanitarian analysts in order to improve humanitarian response. Such data charity practices could entail hidden, exploitative attempts to increase sales (e.g. offering international roaming discounts for cell phone or internet usage to refugees) under the veil of philanthropism. Such practices would not only be restricted to humanitarian action but very commonly might also take place in development aid settings. Because humanitarian action is dependent on development aid when it comes to innovation and experimentation (to ensure informed consent and elements of research ethics and data agency), specifically the principle of independence is of utmost importance in the development field. To prevent the misuse of data charity for commercial goals, humanitarian and development agencies should contractually ensure that when accepting data and technology charity commercial enterprises abstain expressly from financially profiting from the results of development and humanitarian research.

It is important to note, that the humanitarian principles do per so not bind developmental actors. Development aid is very much focused on finding solutions to structural problems such as, indeed, inequalities, corruption or bad governance practices. In essence, by its necessity development aid, in contrast to its apolitical cousin, humanitarian action, is very political. We have, however, also illustrated that the problems that can result from bad data practices are not specific to humanitarian action but concern the developmental context equally. It would, therefore, make sense to extend some of the ideas contained in the extension of the humanitarian principles for data and innovation practices also to the field of development aid. Most importantly, a leading principle should always be

that the use of data should never increase the vulnerability of people to political and economic exploitation or expose them to further risks.

Conclusion

In general, the humanitarian and development communities need to take a much stricter approach to data protection and innovation. Currently, we are experiencing a Wild West attitude regarding the use of data and concerning “experimental techniques” in humanitarian and development settings. This chapter gave an illustrative overview of the current developments and future challenges around data-driven practices in the field of international humanitarian and development aid.

It argued specifically that, while Big Data and data-driven innovation in these two fields of international politics and governance can have tremendous benefits, the risks must not be overlooked. These risks range from the misuse of personally identifiable information (PII) and demographically identifiable information (DII) by criminal and armed political actors, to the increase of vulnerabilities of people affected by disaster, to macro-level shifts of power from (legitimate) political actors towards commercial enterprises. Especially, the latter requires more research in the future as ownership over and distribution of data as new raw material increasingly change the political landscape in the real world and the cyber domain simultaneously.

This chapter also briefly raised developments around the regulation of the use of data-driven innovations and operations and more specifically extended the application of the humanitarian principles to the field of development aid and to the field of Big Data. The demand of the core principle “humanity”, however, remains the same: namely, that aid must not increase the vulnerability of people affected by disasters and crises. Because the spirit of innovation has also animated humanitarian and development agencies towards data-driven innovation, the importance of the role of data protection and deletion policies as well as high standards of research ethics is increasing more than ever before.

While the issues raised do already have a space in the discourse around data innovation and aid, there are emergent issues that have yet to be framed and investigated more concretely for their potential value. This concerns for example the role that blockchain and smart contracts can play in humanitarian and development aid.49 Furthermore, the

---

regulation of states, corporate and other non-state actors in armed conflict recently led to the call for a Digital Geneva Convention.\textsuperscript{50} One of the yet underexplored fields, is the utility of bots (small programmes executing tasks as virtual agents, e.g. chat bots, crawlers, etc.) in open source humanitarian intelligence.\textsuperscript{51}

Big Data and data-driven innovation promise to stay and to further disrupt the way in which international relations in general and humanitarian and development aid in particular are being conducted. It remains crucial that, with all the promises technological innovations hold for the greater good, we do not lose sight of the ethical principles that keep us human. In the end, it boils down to Kant’s second categorical imperative, not to use others as means for innovation towards the greater good, but to see them always also as an end by themselves.\textsuperscript{52}

Bibliography


\textsuperscript{52} Immanuel Kant, \textit{Grundlegung zur Metaphysik der Sitten} (Berlin: L. Heimann, 1870).


Raymond, Nathaniel A. “Beyond ‘Do No Harm’ and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society’s Use of Data.” In Group


