Chapter 5
"Do No Harm" in the Age of Big Data: Data, Ethics, and the Refugees

Patrick Vinck, Phuong N. Pham and Albert Ali Salah

Abstract Leveraging call detail records for humanitarian analysis involves the collection and sharing of a large set of behavioral data, from hundreds of thousands of people. There is a risk that such data could be misused for surveillance and suppression, and there are strong criticisms that have been leveled at efforts involving call detail records. The D4R Challenge is not immune to these criticisms, and during the design and implementation of the challenge, these issues were discussed at length. This chapter outlines these issues and how they were (imperfectly) addressed.

5.1 Introduction

In recent years, the use of information technologies has received considerable attention from the humanitarian community. The interest is not new—humanitarians and scholars have noted since the early 1990s the remarkable transformation brought about by the ability to capture accurate and timely information in real or near-real time, and the possibility of connecting remotely with affected communities using new information and communication technologies [18]. The spread of mobile devices and connectivity, the increased availability of various types of data from nongovernmental sources, and the rise of social media platforms have accelerated this transformation. As a result, assessments and advocacy efforts using information technologies and the accompanying digital data they generate have become a common component of humanitarian work.

P. Vinck · P. N. Pham
Harvard University, Cambridge, MA, USA
e-mail: pvinck@hsph.harvard.edu

P. N. Pham
e-mail: ppham@hsph.harvard.edu

A. A. Salah (✉)
Utrecht University, Utrecht, The Netherlands
e-mail: a.a.salah@uu.nl

Boğaziçi University, Istanbul, Turkey

A. A. Salah et al. (eds.), *Guide to Mobile Data Analytics in Refugee Scenarios*, https://doi.org/10.1007/978-3-030-12554-7_5

The examination of humanitarian digital efforts, however, suggests that the opportunities, limitations, and risks associated with these digital affordances (i.e., actions enabled by technology) have not always been clearly and precisely identified. Specifically, considerable challenges are emerging because of the datafication of humanitarian work—the transformation of virtually all aspects of humanitarian work into quantifiable, machine-readable data, easily manipulated on a computer. These challenges, ethical considerations, and opportunities must be clearly articulated so that solutions can be identified.

As it turns to technology and quantitative tools, the humanitarian community follows other disciplines and a global trend which fundamentally changes how data are produced, managed, analyzed, stored, and utilized [10]. Recent global statistics suggest that three out of four people have a mobile phone and 75% of these mobile phones are smartphones with broadband capabilities and integrated GPS [9]. People use these phones to communicate, post, and view social media content or conduct business, and each of these interactions generates unprecedented amounts of data. At the same time, private companies routinely generate data at a level of precision and granularity that was formerly available only to the intelligence community. Thus, humanitarian work takes place amidst gradual acceptance of sharing personal information with Facebook, Google, Apple, and other tech companies, and increased capabilities to gather, mine, and analyze all kinds of data for surveillance purposes [33]. This data revolution has sparked debates, prominent leaks on surveillance methods [26], and questions around ethical principles and standards, privacy, consent, representativeness, data protection, and data validity and accuracy. Similar concerns exist across disciplines and applications, but because of the risks and high stakes associated with humanitarian work, many practitioners and scholars acutely feel the need to advance the responsible use of data and technologies.

This was a major concern when designing and implementing the Data for Refugees Challenge [24]. As they cross borders, people forcibly displaced are guaranteed a number of rights.[1] In theory, this includes the right to privacy. In practice, however, refugees must provide personal data to numerous government, international, and humanitarian aid agencies as a condition for assistance. They have little to no control over how those data are used and protected. Recently, numerous principles and guidelines have extended the protection of civilians to the protection of their data. These new protection principles note that people should not be put at risk as a result of the way that humanitarian actors record and share information and call for clear and comprehensive data protection policies [2, 28]. In Europe, the General Data Protection Regulation (GDPR) is now in practice, which regulates many of these issues quite strictly. All across European research institutions, training is offered to scholars for GDPR compliance when working with data recorded from humans [34]. Humanitarians, however, are relatively ill-equipped to ensure such dig-

[1] As mentioned in Chap. 1, while Turkey is party to the 1951 Geneva Refugee Convention, it does not grant Syrian refugees the legal status of "refugee," but considers them "temporarily protected foreign individuals". This complicates the rights Syrian refugees have from a legal point of view, where they cannot benefit from the internationally established measures of protection. See Chap. 6 for definitions of the key terms.

ital data protection and the risks are poorly understood. The D4R Challenge is not immune to criticisms that have been leveled at similar efforts leveraging call detail records [29]. There are potential biases, risks to data subjects, and the potential that data and techniques could be used for surveillance and control rather than positive outcomes. These challenges and how they were (imperfectly) addressed are outlined in this chapter.

This chapter started by noting how data science is challenging and transforming humanitarian action. This premise may give the impression that data science and information technologies are merely neutral tools that can benefit or harm humanitarian endeavors depending on how they are used and who uses them. Furthermore, it could be implicitly assumed that used responsibly, ethically, and effectively, data and technologies provide part of the solution to pressing global problems [5], including humanitarian action. Technology itself and the data it generates are seen as neutral facilitators that can be leveraged for social good.

Nothing could be further from the truth, however. Technological mediation, including data analytics, is not neutral. Before their potentially serious environmental, social, and human consequences are even considered [12], data are the results of algorithmic choices and human-designed protocols that have inherent flaws and biases of deep concern for humanitarian action. Data collection and analysis carry significant risks of discrimination and targeting of groups and individuals, potentially resulting in denial of services and basic rights. How data are generated can reflect widespread biases that exist in society [4] and could even exacerbate inequalities [19]. Furthermore, technological actors generating data do not exist in a vacuum devoid of ethical and human rights concerns, which is especially true during humanitarian crisis. Organizations like Palantir Technologies[2] provide data analytics support to humanitarian actors, while simultaneously equipping parties to various conflicts with unique data intelligence capabilities [20]. Cellphone companies may share data with humanitarian organizations, but they also seek to monetize their data and gain market insights. More generally, corporate data practices have raised both ethical and legal issues, particularly concerning the use of personally identifiable information without consent. These, and many more examples, illustrate why data and technology are more than tools, and demonstrate the need for data scientists and technologists to learn and engage with humanitarians just as much as humanitarians and scholars must engage with data science. The D4R Challenge was an opportunity for such engagement. In this chapter, we outline five critical topics that we believe will require the engagement of practitioners and scholars across both the data science and human rights disciplines.

[2] www.palantir.com.

5.2 Adoption of Innovation and Ethical Concerns

Digital affordances have been embraced across disciplines to gather insights into human behaviors. The transformation of humanitarian action along these lines may have been slower, or even met with strong resistance, for several reasons such as a mistrust of corporate actors who own and generate these data, and a traditional emphasis on immediate response and action. Arguably, the lack of clarity on ethical issues concerning the use of various data sources and technologies may also have influenced the rate of adoption of tech-enabled innovation by humanitarians. Everett Rogers' work on diffusion of innovation can help us understand the adoption curve of data innovations among humanitarian organizations [23]. The two extreme categories according to Rogers are (1) the innovators and early adopters on one side—those who adopt technology early, the risk takers and pioneers who lead the way and (2) the late adopters and "laggards", those who wait until they are convinced that the technology works in their best interest and/or resist until necessary. The clarification of ethical principles, risks, and opportunities can greatly influence the behaviors of actors in these categories and should be of broad concern to the human rights movement.

Over the last decade, the early adopters of technology have been found largely outside of traditional humanitarian groups. Individuals and emerging organizations, often grassroots efforts, coalesced and leveraged new technologies to leapfrog traditional humanitarian assessments and actors. Citizen journalism, crowdsourcing, and mapping platforms, for example, have been largely pioneered by new actors. The challenge was that these innovators and early adopters, as Rogers noted, were risk takers. Their concern for guaranteeing the safety of informants, the accuracy, and security of data, and more generally for adhering to ethical and technical standards was largely limited by their lack of experience. Major concerns have been identified and efforts undertaken to establish ethical practices, but these were matched or outnumbered by high-risk efforts and critical failures reflecting the lack of standards and accountability mechanisms, and the absence of organizations mandated to assess risk and develop best practices. A shared, formalized outline of ethical principles and definitions of risk and harm in this context has also been notably absent.

Late adopters, on the other hand, have largely avoided adapting their data practices to the modern era. Methods of inquiry are often dictated by practical realities such as the experience and expertise of the humanitarian organizations, which rarely includes new technology. Among this group, data collection, analysis, and reporting sometimes lack methodological rigor. For them, concerns about ethical principles and standards, privacy, consent, representativeness, data protection, and data validity and accuracy represent principally arguments against using new technologies and the data they generate. This cautious approach is respectable in the high stakes context of human rights data, but it fails to acknowledge the positive examples and benefits of new streams of data. Thus, late or lack of adoption of data innovation may potentially create missed opportunities to improve our understanding of humanitarian crises.

What these two groups have in common is the need to learn how to handle and leverage data, unlocking its value for affected communities while respecting rights

and ethical principles. This includes ensuring the digital protection of already vulnerable populations. This is critical for early adopters to learn to manage risks and for late adopters to recognize the value of new data types.

The formalization of data ethics in the context of modern technologies—such as through enhanced data ethics literacy—may also benefit data scientists and tech companies. Ethics is rarely recognized as an important and relevant consideration in product, service and organizational innovations [1]. Privacy may be the exception because technologies are open to the scrutiny of their users and civil society in general, and recurring scandals where changes in privacy settings have affected some of the most popular new technologies and social networks. Further recognizing and formalizing the centrality of the ethics of data in innovation would likely build trust and help users better understand their risks in using digital platforms. For example, it could help achieve the right balance between protection of and access to personal information, and how that information is used. In that sense, the D4R Challenge offers new insight as to the potential applications of call detail records in humanitarian action and the ethic challenges it raised. Chiefly, as noted above, the notion of informed consent is largely absent when considering CDRs. Users sign a general agreement that let service providers use data with little restriction. However, as noted above, among refugee population, the notion of informed consent is hindered by the link that exists between the sharing of information and access to assistance. A fundamental rethinking of what consent means and how it is obtained may be necessary, but it is clearly beyond what the D4R Challenge could achieve. Rather, the Challenge concentrated on other critical aspects such as responsible data practices and the minimization of risk.[3] It also made sure that a broad set of stakeholders, including refugees and institutions protecting the rights of refugees, participated in the decision processes, through participation in the Project Evaluation Committee (PEC).

5.3 Responsible Data in the Digital Age

Refugees are increasingly tech savvy and exist both in the physical and digital space. Data generated by cellphone and Internet users are on the rise everywhere and will ultimately become nearly ubiquitous. At the same time, governments and perpetrators of human rights violations that forcibly displace millions of people around the world are learning quickly how to leverage public data, networks, and technologies to identify sources of information, spread rumors and fake data, and attempt to use, evade, or adapt to surveillance capabilities. Connectivity and mobile technology are also revolutionizing how smuggling and trafficking of goods and persons take place, especially in conflict settings.

The technologizing and datafication of humanitarian action is a natural response to these changes. However, it also fundamentally changes what it means to be a

[3] See Chap. 1 of this volume.

humanitarian organization. The few prominent successes in using information technologies and data to advance response tend to overshadow the very real challenges in establishing an ethical and responsible data culture. As humanitarian actors become increasingly holders of massive volumes of digital information—or data organizations—their roles and responsibilities toward the protection, sharing and use of the data they collect are evolving. Humanitarian actors take on enormous legal and ethical responsibilities that they are often ill-equipped to handle both in terms of systems and protocols, but also in terms of culture and attitudes toward privacy. The increased responsibility to protect data will be especially challenging. Data held by humanitarian organizations, especially concerning refugees, are almost by definition sensitive, even when they hold no identifiable or personal data or metadata that can be harvested. It can be used to identify individuals or make inferences about groups and communities. In contexts of conflict and other forms of violence, this can be especially sensitive and result in adverse consequences for those whose data have been exposed [6].

Arguably, data held by cellphone service providers are even more sensitive—as they do include detailed personal information and can be used for a wide range of purpose including surveillance. While the D4R Challenge excluded personal information at the design stage of the databases it constructed (i.e., data protection by design and default), it also had to take measures, such as decreasing the granularity of information, to deal with possible adversaries with access to extrinsic and detailed data that could potentially be combined with mobile CDR. It is well known that social media data, for instance, could be very rich in providing personal information. Photographs shared on social media sites with timestamps and GPS locations can easily provide bodies with access to sufficient computing power an accurate way of identifying and tracking millions of people.

Rules around data protection and data sharing are changing, meaning that humanitarian organizations need to be acutely aware of local laws and regulations. For example, several countries now limit the ability of organizations to export data, instead, requiring storage on local systems, especially when identifiable and sensitive data are involved. The EU Data Protection Directive,[4] for example, prohibits personal data from being exported outside the EU or EEA unless appropriate protection is guaranteed.

Humanitarian organizations need to develop appropriate protocols to take into account existing data-related legislation. This kind of legislation adds an additional layer of regulatory compliance, on top of the already-present difficulty of ensuring privacy when records are at risk of being subpoenaed by national justice systems. The D4R had to work within a strict legal framework and review by the authorities, guided by the legal team of Türk Telekom. This task was much facilitated by the legal precedent of the previous Data for Development Challenges, as described in Chap. 1 of this volume.

How humanitarians communicate and use data are also critical. In the past, reports published and discussed in a major Western hub had little chance of being scruti-

[4] https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en.

nized in remote corners of the world. This is no longer the case. This brings positive opportunities for individuals, groups, or communities to speak for themselves. However, it also means that perpetrators of human rights violations are more aware of what has been reported and may act, for example, by targeting possible sources of information. Humanitarians must now learn new ways to control and manage the information space in which they operate.

The challenges outlined above call for a closer collaboration and transfer of knowledge and experience with established data and technology companies. The D4R Challenge opens prospects for such collaborations. Tech companies, including cellphone service providers, however, are not immune to these risks, as illustrated by the recent hack of Yahoo's three billion email accounts[5] or the hacking of credit score company Equifax.[6] In this new reality, clear data protection plans and articulation of responsibilities are needed, including responsibilities toward those whose data may be compromised.

Finally, the responsible use of data should also entail having the capability to understand data limitations and biases, and to leverage multiple streams of data for analysis. One common argument in favor of sensor data (generated passively, without user knowledge) is that because the data result from user behaviors in their natural environments, without "observers," it would avoid biases from the artificial condition of having researchers present [15]. However, there is ample evidence of online lies, manipulation of behaviors, and purposeful misinformation. Enhanced data literacy and ground-truth annotation are needed in order to verify sensor data. What should be avoided is a sterile debate on what type of data is superior; exploring how multiple data streams can be best used to generate unique insights is a far more productive avenue.

5.4 Ethics as a Common Standard Across Organizations

In addition to gathering their own data, some humanitarian organizations opt to enter into partnerships in order to obtain data from satellites, cell phone networks, online platforms, and other data sources. Opportunities for such partnership are at the core of the initiatives like the D4R Challenge. Most of these kinds of data are generated, collected and processed under the auspices of private-sector corporations [30]. Humanitarians need to position themselves in relation to the complex issues raised by entering into this kind of public–private partnerships, and possibly advocate for improved ethical standards.

Furthermore, the political climate under which the data are being shared could be a source of bias itself. The extended control of corporations on the publication of results based on the data is both motivated by the need to ensure that the data subjects are not harmed in any way, but also by the need to make sure that the

[5] https://www.nytimes.com/2017/10/03/technology/yahoo-hack-3-billion-users.html.

[6] https://www.ftc.gov/equifax-data-breach.

results do not endanger the corporations' relations with the power holders. This is a particularly sensitive issue for countries where legal institutions and free press do not function well, which puts the corporations on a precarious edge. For refugee-related issues, a lot of responsibility falls on governmental institutions, whose functioning and efficiency can be realistically assessed with the processing of rich data sources. This is obviously both a blessing and a curse.

At the moment, how, when, and under what conditions corporate actors share data is at best ad hoc and lacks transparency or coordination. Companies like DigitalGlobe have publicly partnered and released imagery during crises,[7] but other data holders, especially those with commercial interests on the ground have had less incentive and willingness to do so, or have shown the willingness to "change the rules" once organizations become dependent on their data, demanding ever-increasing payment to access data.

Collaborating across organizations with significantly different data cultures is challenging. Actions that may be perceived as acceptable in one company—say, for example, the release of call detail records during a crisis—may in fact be unethical or even illegal [17]. How intellectual property is assigned may also be perceived very differently across organizations. Yet this type of collaboration is almost unavoidable in today's humanitarian sector. Again, this issue highlights the importance of formalizing data ethics in the context of modern technologies and the need to develop data ethics literacy across organizations. It also requires exploring the role and responsibilities that companies have in enabling or protecting the free exchange of ideas.

5.5 Embracing Ethical Complexity and Emerging Rights

Research at the nexus of data science and humanitarian action is largely focused on how data are used and the conditions under which responsible data practices can transform humanitarian efforts. In this context, a common argument is that the application of existing ethical principles grounded in the recognition of the dignity of the person must guide the responsible use of data for humanitarian action [22]. These include expectations of informed consent, voluntary participation, the well-being and security of participants, and balancing risks and benefits [32]. Yet, in most cases, data are being created, collected, mined, analyzed, monitored, sold, stored, and used for diverse reasons, mostly beyond individuals' control of the data generated about them [11].

As we have noted elsewhere [14], the discussion of ethical principles, dilemmas, and risks in collecting and sharing CDRs must build on several decades of progress in understanding and defining principles for ethical research. Similar principles have historically been developed primarily in the biomedical and behavioral sciences. The practice of Big Data analytics, and specifically the use of CDRs, closely resembles research cycles and processes, and the insights sought are relevant to behavioral

[7] https://www.digitalglobe.com/ecosystem/open-data.

science. While arguably corporations are not research institutions, lessons can be learned and modeled from these more developed ethical frameworks and applied to these new emerging fields. There are a number of landmark guides for ethical research principles as laid out in the Nuremberg Code,[8] Declaration of Helsinki,[9] and Belmont Report.[10] A more recent initiative by the US Department of Homeland Security, Science and Technology, Cyber Security Division revised, and adapted established ethical principles in the context of the ICT and data revolutions. The result was published as The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research.[11]

Less attention has been paid to the human rights consequences of data collection practices, for example in undermining the right to self-expression and freedom of association. A fundamental question is whether humanitarian action in general undermines the rights to privacy and to autonomy in decisions concerning a person's own welfare, and at the most extreme, possibly feeds a regime of continuous surveillance [3] and hyper-targeting. Under such a regime, data are gathered with the specific objective of influencing actions in a way that may not be beneficial to the individual whose actions are being influenced, in other words, in a manner that consciously seeks to undermine autonomy or to discriminate. There is a major risk that information collected from refugees will be publicly available for a much longer duration than intended. Perhaps decades after they have resettled, troves of refugee data may be used to affect the services that those formerly displaced are provided with. The right to be forgotten, as defined by the European Commissioner for Justice, Fundamental Rights, and Citizenship, directly addressed this issue, seeking to protect informational self-determination, or autonomy [31].

More generally, the effects of poor data collection and management practices are largely considered to be in the form of physical violence, retribution, or shaming [13]. The potential for more complex and far-reaching impact is not well understood, including, for example, risks related to the capacity to re-identify data because of advances in computing and communications technology. Understanding risks, however, is an ethical imperative. To give a simple, but illustrative example: Mandating the storage of data in a secure server through the user agreement is a common precaution, but ignores the fact that the security afforded by systems today may be easily bypassed by the technology of tomorrow. The D4R Challenge mandates complete destruction of the data at the end of the project term to deal with such a potential breach (see Chap. 1, Appendix).

Emerging explorations of data ethics in humanitarian action also focus on data ownership and the limits of what can be done with information collected among affected people with limited abilities to provide fair and informed consent. The need to create and adopt protection and privacy standards has emerged and broad efforts at

[8] https://history.nih.gov/research/downloads/nuremberg.pdf.

[9] https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/.

[10] https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/.

[11] https://www.dhs.gov/publication/menlo-report.

establishing a digital do no harm policy are being actively pursued [6]. These efforts can be likened to a digital equivalent of Geneva Conventions, which seek to protect, in war and conflict situations, people not taking part in hostilities and those who are no longer doing so.

The idea of a set of Digital Geneva Conventions has been most famously championed by Microsoft [27]. The firm's products have been the target of nation-state attacks, prompting the call to ensure the protection of corporate assets and civilians' data in times of war and other situations of violence, with the active assistance of technology companies. The proposition recognizes the need to expand the do no harm framework to "critically assess" how using new technologies can potentially expose already vulnerable populations to further risks and insecurities, even where intentions are at their best and conditions at their most challenging [6]. Beyond protection, new approaches are needed to ensure data agency and ownership at the individual and community level, and individual and collective mechanisms for redress and restitution for digital harm.

Additional complexity emerges from the ability to combine and recombine data in so many ways that they enable making inferences about groups. This is critical because individual data are no longer only useful for gaining information about and targeting the individual, but also—and perhaps above all—for gaining information about and targeting groups. In that sense, the mere fact of being associated with a group—even if no data were ever shared – provides insight about individuals who never shared data. This challenges the idea that human rights must be borne by individual humans and therefore do not apply to groups [7]. Indeed even when group membership is central to human rights (e.g., prohibition of discrimination or persecution), the right itself is held by individuals—not the group itself [8]. For example, the extraction of DNA may pose a significant risk to the privacy rights of individuals from whom the sample was extracted, but also to their related genetic group. Close relatives and broader groups they may belong to (e.g., ethnic group) must also therefore have their interests taken into account. Creating a group privacy right, or a right to be forgotten, might provide effective protection, but only if it can be enforced. The definition of new and emerging rights and ethical principles, and their enforcement, should be at the center of the engagement between human rights scholars, practitioners, and data scientists.

5.6 Linking Data to Action

Information technologies contribute to the long humanitarian tradition of building strong narratives and visuals to generate attention. Most commonly, the documentation of humanitarian needs seeks to identify who is affected, how and how much they are affected, why they are affected, and what to do about it. Satellite imagery, for examples, has produced before and after images of human rights violations and served to document trends and ongoing actions. Whether or not these data impact actions, policies, and intervention on the ground, preventing mass violence or geno-

cide, as is sometimes claimed, is open to debate [21]. There is a lack of evidence of any protective or preventative effect [25]. Some incidents even suggest that the use of ICT's by humanitarian and human rights organizations led to negative outcomes [16].

In an increasingly connected world, collecting data and featuring analysis and results without a link to action, however, becomes an increasingly difficult proposition. Those who hold actionable data have a moral obligation and duty to take action. Organizations who have coalesced around information technologies as a service for humanitarians may argue that they are merely a platform that hosts information. But the response gap—the difference between the needed response and what actually occurs—is widening as the humanitarian community becomes increasingly apt at documenting what is happening, where and to whom in near-real time. That progress is not accompanied by a similar improvement in understanding the root causes of what is happening and what to do about it. The amount of data now available undermines any notion that the events were unknown, but it also raises the level of expectation that people will be saved, that actions will be taken. The failure to do so significantly undermines humanitarian efforts and may even create resentments among those who did not receive the help or support they needed.

5.7 Conclusion

This chapter discussed five critical topics, far from an exhaustive listing of all ethical challenges emerging at the nexus of humanitarian action and data science. Rather, it outlined a possible multi-faceted research agenda that will benefit not only humanitarian organizations, but also data science. Ethics, we argued, is a critical issue in the diffusion of data innovation, yet it is rarely recognized as an important and relevant consideration in product, service and organizational innovations. The absence of a shared, formalized outline of ethical principles and definitions of risk and harm means that early adopters may underestimate risks, while late adopters overestimate them.

Critically, humanitarian organizations are increasingly becoming data organizations (see for instance the next chapter of this volume), a move for which they are not always prepared, including in their interaction with emerging corporate and grassroots actors. Research institutions can assist such organizations by providing new tools; governments and corporations can provide timely data. The transformation of humanitarian action as a result of the data revolution, however, goes further than merely providing new data and tools. It is raising fundamental questions about the possibility of some rights (e.g., autonomy) and the emergence of new ones (e.g., group privacy). Re-defining ethical principles and their relations to human rights and data sciences is a broad proposition, but it is critical to the relevance and integrity of human rights work. The D4R Challenge was not meant to solve these problems. Rather, it offers new insights into how these challenges concretely play out when exploiting privately held data for humanitarian purposes.

References

1. Brusoni S, Vaccaro A (2017) Ethics, technology and organizational innovation. J Bus Ethics 143(2):223–226
2. CHS (2014) The core humanitarian standard on quality and accountability. CHS Alliance, Group URD and the Sphere Project
3. Couldry N (2017) Surveillance-democracy. J Inf Technol Polit 14(2):182–188
4. Crawford K, Schultz J (2014) Big data and due process: toward a framework to redress predictive privacy harms. BCL Rev 55:93
5. Cukier K, Mayer-Schoenberger V (2013) The rise of big data: how it's changing the way we think about the world. Foreign Aff 92:28
6. Jacobsen KL (2015) Humanitarian technology: revisiting the do no harm debate. ODI humanitarian practice network. https://odihpn.org/blog/humanitarian-technology-revisiting-the-%C2%91do-no-harm%C2%92-debate/
7. Jones P (2017) Human rights, group rights, and peoples' rights. In: Human rights. Routledge, pp 277–304
8. Kammourieh L, Baar T, Berens J, Letouzé E, Manske J, Palmer J, Sangokoya D, Vinck P (2017) Group privacy in the age of big data. In: Group privacy. Springer, pp 37–66
9. Kemp S (2017) Digital in 2017 global overview report. We are social and hootsuite. https://wearesocial.com/special-reports/digital-in-2017-global-overview/
10. Kitchin R (2014) The data revolution: big data, open data, data infrastructures and their consequences. Sage Publications
11. Kosciejew M (2014) Proposing a charter of personal data rights. Inf Manag J 48(3):27–32
12. Kranzberg M (1986) Technology and history: "Kranzberg's laws". Technol Cult 27(3):544–560
13. Latonero M, Gold Z (2015) Data, human rights & human security. Hum Rights Hum Secur
14. Letouzé E, Vinck P, Kammourieh L (2015) The law, politics and ethics of cell phone data analytics. Data-pop alliance white paper series data-pop alliance, World Bank Group, Harvard humanitarian initiative, MIT Media Lab and Overseas Development Institute
15. Levitt SD, List JA (2007) What do laboratory experiments measuring social preferences reveal about the real world? J Econ Perspect 21(2):153–174
16. Mancini F, Letouze EF, Meier P, Vinck P, Musila GM, Muggah R, Diniz G, Puig Larrauri H, Matveeva A, O'Reilly M (2013) New technology and the prevention of violence and conflict. United States Institute of Peace
17. McDonald SM (2016) Ebola: a big data disaster-privacy, property, and the law of disaster experimentation. The Centre for Internet and Society
18. Meier P (2011) New information technologies and their impact on the humanitarian sector. Int Rev Red Cross 93(884):1239–1263
19. O'Neil C (2017) Weapons of math destruction: how big data increases inequality and threatens democracy. Broadway Books
20. Parker B (2019) New UN deal with data mining firm Palantir raises protection concerns. The new humanitarian. http://www.thenewhumanitarian.org/news/2019/02/05/un-palantir-deal-data-mining-protection-concerns-wfp/
21. Parks L (2009) Digging into Google earth: an analysis of "crisis in Darfur". Geoforum 40(4):535–545
22. Pham PN, Vinck P (2012) Technology, conflict early warning systems, public health, and human rights. Health Hum Rights 14(2):106–117
23. Rogers EM (2003) Diffusion of innovations, 5th edn. Simon & Schuster International
24. Salah AA, Pentland A, Lepri B, Letouzé E, Vinck P, de Montjoye YA, Dong X, Dağdelen Ö (2018) Data for refugees: the D4R challenge on mobility of Syrian refugees in Turkey. arXiv preprint arXiv:180700523
25. Sandvik K, Raymond N (2017) Beyond the protective effect: towards a theory of harm for information communication technologies in mass atrocity response. Genocide Stud Prev Int J 11(1)

26. Scheuerman WE (2014) Whistleblowing as civil disobedience: the case of Edward Snowden. Philos Social Crit 40(7):609–628
27. Smith B (2017) The need for a digital Geneva convention. The official Microsoft blog, 14
28. Sphere (2018) The sphere handbook: humanitarian charter and minimum standards in humanitarian response, 4th edn. Practical Action Publishing
29. Taylor L (2016) No place to hide? The ethics and analytics of tracking mobility using mobile phone data. Environ Plann D Soc Space 34(2):319–336
30. Taylor L, Broeders D (2015) In the name of development: power, profit and the datafication of the global south. Geoforum 64:229–237
31. de Terwangne C (2014) The right to be forgotten and informational autonomy in the digital environment. In: The ethics of memory in a digital age. Springer, pp 82–101
32. US Department of Health, Education, and Welfare (1978) The national commission for the protection of human subjects of biomedical and behavioral research, The Belmont report: ethical principles and guidelines for the protection of human subjects of research. Department of health, education, and welfare, U.S
33. Van Dijck J (2014) Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. Surveill Soc 12(2):197–208
34. Voigt P, von dem Bussche A (2017) The EU general data protection regulation (GDPR). A practical guide, 1st edn. Springer International Publishing, Cham