



Selected Papers of #AoIR2020:
The 21st Annual Conference of the
Association of Internet Researchers
Virtual Event / 27-31 October 2020

NON-HUMAN HUMANITARIANISM: WHEN AI FOR GOOD TURNS OUT TO BE BAD

Mirca Madianou
Goldsmiths, University of London

With over 168 people needing humanitarian assistance in 2018 and over 69 million refugees, the humanitarian sector is facing significant challenges. Proposals that artificial intelligence (AI) applications can be a potential solution for the crises of humanitarianism have been met with much enthusiasm. This is part of the broad trend of 'AI for social good' as well as the wider developments in 'digital humanitarianism', which refers here to the uses of digital innovation and data within public and private sectors in response to humanitarian emergencies. Chatbots; predictive analytics and modeling that claims to forecast future epidemics or population flows; and biometric technologies, which rely on advanced neural networks which employ machine learning algorithms, are some of the examples which are becoming increasingly popular in aid operations.

The paper develops an interdisciplinary framework that brings together colonial and decolonial theory, the critical inquiry of humanitarianism and development, critical algorithm studies as well as a sociotechnical understanding of AI. Humanitarianism here is understood as a complex phenomenon: not just the 'imperative to reduce suffering', as it is usually defined (Calhoun, 2008), but also as an industry, a discourse, and a historical phenomenon with roots in 19th and 20th century colonialisms (Fassin, 2012; Lester & Dussart, 2014). AI is an equally multifaceted phenomenon: not just a technological innovation based on advanced computation and machine learning algorithms, but also an industry as well as a particular discourse about technology. AI can only be understood together with data and algorithms – the three are inseparable as AI depends on machine learning algorithms which are the product of particular datasets. Given 'big data' are inherently incomplete and have ontological and epistemological limitations (Crawford & Finn, 2014), AI applications reproduce and potentially amplify existing biases found in large datasets (Benjamin, 2019; Eubanks, 2018; Noble, 2018 among others).

Suggested Citation (APA): Madianou, M. (2020, October). *Non-human humanitarianism: when AI for good turns out to be bad*. Paper presented at AoIR 2020: The 21th Annual Conference of the Association of Internet Researchers. Virtual Event: AoIR. Retrieved from <http://spir.aoir.org>.

Empirically the paper is based on a review of key AI applications in the humanitarian field. In particular, my analysis will focus on chatbots, predictive analytics and modelling.¹ Apart from analysing the actual innovations (eg, the functionality of chatbots; modeling outcomes, data visualisations) the paper also draws on interviews with seven groups of key stakeholders, from humanitarian officers to entrepreneurs and digital volunteers as part of a larger study of digital humanitarianism (Madianou, 2019a).

The analysis suggests that some of the AI developments do not fulfill grand claims such as interactivity, let alone 'intelligence'. For example, chatbots, which are computer programmes designed to interact with humans online as though they were a person, have been developed by a number of aid organisations in order to improve information dissemination and collect feedback from affected communities. My analysis reveals that some of the most prominently advertised chatbots are hardly interactive. In fact, some of the chatbot functionalities could have been easily replaced by SMS messages, or even analogue technologies such as leaflets. Chatbots didn't provide added value, for example through an opportunity to answer questions beyond the template of a very limited number of options. Furthermore, predictive modeling programmes often appear to summarise information that is already available in the public domain. Other applications, such as those that analyse refugee Call Detail Records (CDR) made available by Mobile Network Operators in order to estimate refugee integration in host societies, produce findings that can be captured by methods that entail fewer risks to the research subjects. CDRs are extremely sensitive metadata especially when linked to already vulnerable groups such as displaced people.

The observation that some of the key developments in AI humanitarianism fail their own objectives should not mean that these innovations do not have powerful consequences. First of all, some of the applications analysed in the paper entail very high risks for vulnerable groups with few safeguards against potential data breaches, which are increasingly common in the humanitarian sector. More broadly, 'predicting' or 'forecasting' is beyond the humanitarian remit, which involves responding to emergencies. Predicting crises (for example, future refugee flows) is inevitably political which is at odds with humanitarian principles of neutrality.

Automation reproduces human biases found in datasets that train machine learning algorithms. While humanitarian analytics programmes present themselves as infallible and objective, this is far from true. Deferring decisions to processes of automation carries risks of further disadvantaging already marginalised people. We observe a wider shift in the nature of humanitarian work. Digital technology separates actors (humanitarian workers) from the consequences of their interventions. More broadly, the shift to AI and other digital innovations requires humanitarian organisations to outsource these activities to private vendors (McDonald, 2019) turning aid agencies into managers of contracts, rather than providers of aid.

¹ Biometrics, the other large area based on AI developments, is examined elsewhere (Madianou, 2019b) in order to explore in depth the nature of biometric measurements and the ways they connect with the literatures on bodies / embodiment and surveillance, control and securitization.

Despite the limitations of AI interventions, such developments are cloaked in a discourse of inherent progress exemplified in the phrase 'AI for good'. Digital mediation has a long history of erasing its own work (Bolter & Grusin, 2000; Eisenlohr, 2011) which is echoed in the framing of AI interventions as a form of 'magic' that miraculously projects 'the truth'. Data visualisations play an important part in this 'magic-making' process by doing persuasive work (Kennedy et al., 2016). By erasing the work of mediation and cloaking themselves in an aura of 'magic', AI interventions occlude the ways in which they construct realities and crucially, the way they conceal the power relations that sustain the humanitarian system. In so doing, AI interventions reconfirm the hierarchy between 'problem solvers' and 'problem owners' – to draw on the language used in industry events such as the United Nation's Global Summit 'AI for Social Good'. Rather than democratizing the relationships between humanitarian providers and suffering subjects, digital humanitarianism reaffirms the power asymmetries first established in humanitarianism's colonial iteration.

Humanitarian AI appears to solve problems, whilst in practice benefitting stakeholders, including commercial companies which are increasingly involved in public-private partnerships. There is a strong element of experimentation with untested technologies and the data they produce in what are vulnerable regions with little or no regulation for privacy and data protection. The hype generated by humanitarian innovation can benefit the commercial applications of a particular technology. Ultimately, the paper argues that by turning complex political problems like displacement and hunger into problems with technical solutions, AI depoliticizes humanitarian emergencies.

This is not a call for a return to an earlier, purer form of 'analogue' humanitarianism. The non-human aspects of AI humanitarianism reveal, rework and amplify existing deficiencies of humanitarianism. As our analysis reveals that 'AI for social good' can be bad, we conclude with a reflection on the notion of 'good'.

References

Benjamin, R. (2019). *Race After Technology*. Cambridge, UK: Polity.

Bolter, J. D. & Grusin, R. (2000). *Remediation: Understanding New Media*. Cambridge, MA: MIT Press.

Calhoun, C. (2008) 'The imperative to reduce suffering'. In Barnett, M. & Weiss, T. G. (Eds.) *Humanitarianism in Question*. Ithaca, NY.: Cornell University Press

Crawford, K., & Finn, M. (2015). The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal*, 80(4), 491-502

Eisenlohr, P. (2011) Introduction: what is a medium? Theologies, technologies and aspirations. *Social Anthropology*, 19(1), 1-5.

Eubanks, V. (2018). *Automating Inequality. How High-tech tools Profile, Police, and Punish the Poor*. New York: St Martins Press.

Fassin, D. (2012) *Humanitarian Reason: A moral history of the present*. Berkeley, CA.: University of California Press.

Kennedy, H., Hill, R. L.; Aiello, G. & Allen, W. (2016). The work that visualisation conventions do. *Information, Communication & Society*, 19(6), 715-735, DOI: [10.1080/1369118X.2016.1153126](https://doi.org/10.1080/1369118X.2016.1153126)

Lester, A., & Dussart, F. (2014). *Colonization and the Origins of Humanitarian Governance*. Cambridge: Cambridge University Press.

Madianou, M. (2019). Technocolonialism: digital innovation and data practices in the humanitarian response to refugee crises. *Social Media and Society*, vol. 5(3) <https://doi.org/10.1177/2056305119863146>

Madianou, M. (2019b) The Biometric Assemblage: surveillance, experimentation, profit and the measuring of refugee bodies. *Television and New Media*, vol. 20(6): 581-599. <https://doi.org/10.1177/1527476419857682>

McDonald, S. (2019) From space to supply chain: humanitarian data governance. (August 12, 2019). Available at SSRN: <https://ssrn.com/abstract=3436179> or <http://dx.doi.org/10.2139/ssrn.3436179>

Noble, S. (2018). *Algorithms of Oppression: How search engines reproduce racism*. New York: New York University Press.