



Information, Communication & Society

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/rics20>

Nonhuman humanitarianism: when 'AI for good' can be harmful

Mirca Madianou

To cite this article: Mirca Madianou (2021) Nonhuman humanitarianism: when 'AI for good' can be harmful, *Information, Communication & Society*, 24:6, 850-868, DOI: [10.1080/1369118X.2021.1909100](https://doi.org/10.1080/1369118X.2021.1909100)

To link to this article: <https://doi.org/10.1080/1369118X.2021.1909100>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 08 Apr 2021.



Submit your article to this journal [↗](#)



Article views: 392



View related articles [↗](#)



View Crossmark data [↗](#)

Nonhuman humanitarianism: when ‘AI for good’ can be harmful

Mirca Madianou

Department of Media, Communications and Cultural Studies, Goldsmiths – University of London, London, UK

ABSTRACT

Artificial intelligence (AI) applications have been introduced in humanitarian operations in order to help with the significant challenges the sector is facing. This article focuses on chatbots which have been proposed as an efficient method to improve communication with, and accountability to affected communities. Chatbots, together with other humanitarian AI applications such as biometrics, satellite imaging, predictive modelling and data visualisations, are often understood as part of the wider phenomenon of ‘AI for social good’. The article develops a decolonial critique of humanitarianism and critical algorithm studies which focuses on the power asymmetries underpinning both humanitarianism and AI. The article asks whether chatbots, as exemplars of ‘AI for good’, reproduce inequalities in the global context. Drawing on a mixed methods study that includes interviews with seven groups of stakeholders, the analysis observes that humanitarian chatbots do not fulfil claims such as ‘intelligence’. Yet AI applications still have powerful consequences. Apart from the risks associated with misinformation and data safeguarding, chatbots reduce communication to its barest instrumental forms which creates disconnects between affected communities and aid agencies. This disconnect is compounded by the extraction of value from data and experimentation with untested technologies. By reflecting the values of their designers and by asserting Eurocentric values in their programmed interactions, chatbots reproduce the coloniality of power. The article concludes that ‘AI for good’ is an ‘enchantment of technology’ that reworks the colonial legacies of humanitarianism whilst also occluding the power dynamics at play.

ARTICLE HISTORY

Received 21 December 2020
Accepted 22 March 2021

KEYWORDS

Artificial intelligence (AI); ‘AI for social good’ (AI4SG); automation; chatbots; humanitarianism; decolonial theory; refugees

In March 2016, the Silicon Valley start-up X2AI launched ‘Karim’, a psychotherapy chatbot, to support Syrian refugees in Lebanon. The chatbot uses natural language processing, a form of artificial intelligence (AI), to simulate human conversations in Arabic through existing communication channels such as SMS texts or Facebook Messenger. Given the prevalence of mental health issues in situations of displacement and war

CONTACT Mirca Madianou  m.madianou@gold.ac.uk

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

(Charlson et al., 2019) and given the parallel lack of trained psychotherapists in such settings, chatbots have been proposed as a potential solution to meet the demand for help. This particular chatbot was reportedly piloted on 60 Syrians ‘mostly men and boys’¹ – a small pilot for scaling up to a large and vulnerable population: there are over one million Syrian refugees in Lebanon. X2AI developed the pilot in partnership with ‘Field Innovation Team’, a non-profit specialising in technology in disaster recovery, and the so-called ‘Singularity University’, the Silicon Valley business incubator and consultancy service.

Chatbots are increasingly prevalent in humanitarian settings and have been developed both by humanitarian agencies as well as private companies, such as X2AI, as part of ‘Artificial Intelligence for social good’ (‘AI for good’ or ‘AI4SG’) initiatives. AI4SG, which appears to be the prevailing term, refers to the applications powered by computation (such as machine learning) which claim to be addressing social challenges. AI4SG is part of the tradition of ‘technology for good’ and Information Communication Technologies for Development (ICT4D) which include applications in humanitarian or international development settings that typically aim to address the United Nation’s Sustainable Development Goals.² There are various AI applications in humanitarian settings, such as satellite imagery, data visualisations, chatbots and biometrics.³ The United Nations hosts an annual ‘AI for Good Summit’, while most humanitarian agencies include ‘AI for good’ projects, often in partnerships with the private sector.⁴

The article analyses chatbots as emblematic of humanitarian AI applications in order to question some of the assumptions behind the phenomenon of ‘AI for good’. Acknowledging that both humanitarianism and technology are steeped in relations of power the article asks whether chatbots, as exemplars of ‘AI for good’, reproduce inequalities in the global context. Drawing on a wider study of digital innovation, data and automation in humanitarianism which includes interviews, participant observation and digital ethnography, I observe that chatbots do not fulfil the promises of ‘intelligence’. This does not imply that chatbots have no significant consequences. On the contrary, the article argues that chatbots can present significant risks regarding data safeguards and misinformation which can have adverse consequences for some of the world’s most vulnerable people. By reducing participation to a ‘box-ticking exercise’, by extracting value from data and experimentation with new technologies and by asserting Eurocentric values in humanitarian contexts chatbots reproduce the asymmetries between affected people and humanitarian organisations. The article concludes that ‘AI for good’ is an ‘enchantment of technology’ (Gell, 1992) that reworks the colonial legacies of humanitarianism whilst also occluding the power dynamics at play.

The article develops an interdisciplinary framework that brings together the decolonial critique of humanitarianism and critical algorithm studies. This framework is presented in the following section before discussing the empirical context and findings.

A decolonial approach to humanitarianism and AI

The challenges for the humanitarian sector and the promise of technology

Humanitarianism is a complex phenomenon: not just the ‘imperative to reduce suffering’ (Calhoun, 2008), but also an industry (Krause, 2014), a discourse (Chouliaraki, 2013) and a historical phenomenon with roots in nineteenth and twentieth century colonialisms

(Lester & Dussart, 2014). To understand humanitarianism, one has to understand the power relations between mostly western ‘saviours’ and the suffering ‘others’ who are typically former colonial subjects in the so-called ‘global South’. This fundamental asymmetry is at the heart of every humanitarian intervention including those involving ‘AI for good’. Decolonial theory can help explain how humanitarianism reproduces Eurocentric systems of knowledge and social and racial hierarchies which sustain the ‘coloniality of power’, a notion developed by Quijano (2000) to explain ‘western’ dominance globally after the end of colonial rule.

With over 206.4 million people needing humanitarian assistance in 2018, and over 79.5 million forcibly displaced people by the end of 2019, the humanitarian sector faces significant challenges.⁵ The scale and scope of humanitarian operations have grown significantly over the past couple of decades while the sector is marked by an increasing marketisation with intense competition for funding (Krause, 2014). Donors demand efficiencies and evidence of impact, which often explain the push for digital technologies which are celebrated for their cost-cutting potential and for providing metrics for audit (Madianou et al., 2016). The private sector has a very active presence in the humanitarian field, mainly in the form of private-public partnerships, or through private foundations, or corporate social responsibility policies (McDonald, 2019). One such example is the Karim chatbot which was developed by the X2AI Foundation – the non-profit arm of X2AI. Private companies are driven by a wider capitalist logic that reframes political problems as problems with business solutions (Madianou, 2019a).

Digital innovation is proposed as a solution to several of the sector challenges. This includes the pressures to democratise humanitarianism: to increase the participation of, and accountability to affected communities through initiatives such as ‘Communicating with Communities’ (CwC) or ‘Accountability with Affected People’ (AAP) (Madianou et al., 2016). The interactive nature of digital technologies is rendered as an opportunity to facilitate the participation of affected communities in their own recovery and correct the asymmetries on which humanitarianism is based.

Chatbots are the latest iteration in a long lineage of technologies developed to improve information dissemination among and feedback by affected people. Yet, the data trails of some of those earlier applications reveal that the main purpose of feedback channels is to produce audit trails which are in turn used to justify the funding of humanitarian projects. An earlier ethnography of the recovery from Typhoon Haiyan found that the SMS hotlines which were meant to improve feedback, and therefore accountability to affected people, were actually primarily used for audit reporting. The capacity of digital platforms to aggregate data and forward them as evidence of impact played a vital role in this process (Madianou et al., 2016). Apart from the extraction of data, innovation practices also allow the extraction of value from experimentation with untested technologies (Madianou, 2019a). The hype generated by innovation in humanitarian settings translates into heightened visibility and interest in new products and services which is particularly attractive for companies seeking branding opportunities.

AI and ‘AI for good’: mythology and colonial legacies

AI refers to several phenomena including computational approaches (for example, machine learning, which in turn can include advanced neural networks, or natural

language processing); and practices, such as design or coding. AI is also a 'planetary infrastructure' (Crawford & Joler, 2018). Crucially, AI is also a set of discourses around technology. Narratives about AI influence the design of new applications (Natale, 2019), as well as their public perception and regulation. Given the opacity around the workings of algorithms and automation (Pasquale, 2015) the public's understanding of AI is inevitably shaped by popular culture narratives (Craig et al., 2018). The term 'artificial intelligence' is loaded with meaning, implying a 'thinking' machine, even in cases where there is little evidence of autonomous thinking. Anthropomorphism, the attributing of human characteristics such as intelligence to software programmes (for example, when asserting that 'the machine *learns*') carries powerful connotations of agency. At the same time, the use of the term 'nonhuman' to refer to automation and AI, also performs important discursive work. 'Nonhuman' implies a distancing, if not a complete removal of human subjectivity from processes of computation concealing the fact that AI is the result of human labour often in exploitative conditions (Anwar & Graham, 2020).

The above discourses are exemplified in the case of chatbots which are software programmes that can recognise text, or voice-based inputs and interact with humans online (Gehl & Bakardjieva, 2017). Chatbots have a long history going back to ELIZA, the first chatbot created by Weizenbaum (1966) at the Massachusetts Institute of Technology in the 1960s. The idea of the 'thinking machine' can be traced back to Alan Turing, who in 1950 argued that computers were capable 'intelligent behaviour'. The 'Turing test', originally known as the 'Imitation Game', aims to prove whether 'machines can think' in the context of conversations (Turing, 1950). To pass the test, a machine has to produce humanlike responses so that an independent evaluator cannot tell the machine and the human apart.

Not all chatbots would pass the Turing test. The degree of intelligence in AI operations varies dramatically (Broussard, 2018). Low-level chatbots, probably the majority of chatbots, provide answers from a list of predetermined questions. Such chatbots can be useful in handling 'frequently asked questions' where questions and answers are pre-scripted, but they can't respond to more complex questions nor hold a conversation (Bakardjieva, 2015). Advanced chatbots on the other hand, run on natural language processing and machine learning algorithms which aim to imitate human language. Most chatbots are somewhere in the middle: for example, they can respond to keywords or simple phrases, but they are not able to reproduce 'natural language'.

Just as big data are surrounded by 'the aura of truth, objectivity, and accuracy' (boyd & Crawford, 2012), AI is underpinned by a similar 'mythology': a set of beliefs that its 'intelligence' is scientifically proven and therefore more advanced than previous forms of computation or innovation. This is typical of a teleological understanding of technology, where the latest iteration is claimed to be better and more advanced than the previous ones (Kember & Zylinska, 2012). Nowhere is the mythology of 'artificial intelligence' stronger than in the AI4SG phenomenon. The 'for-good' element of the phrase, makes a moral claim about the purpose of technology whilst occluding the underlying power relations. By foregrounding technology as 'good', AI4SG forecloses critical questions regarding the aforementioned asymmetries of humanitarianism and the way they shape humanitarian innovation. Likewise, the emphasis on 'good' obscures the way technology companies often use AI4SG projects to entrench their position in public life. Crucially, the term AI4SG assumes consensus on what 'social good' is, when the term is fundamentally contested. What is good for one group, may not be good for others.

Surveys of AI4SG projects confirm that there is no agreement about what good means (Floridi et al., 2020; Shi et al., 2020). If some consensus exists, it revolves around the orientation towards the UN's sustainable development goals (SDGs) which are the focus of the annual 'AI for Good' Summit.⁶ By closely aligning to the framework of international development, AI4SG inherits some the criticisms regarding development and humanitarianism discussed earlier, namely the reproduction of asymmetrical relationships between the so-called Global North and South. AI4SG projects which take a top-down approach that presupposes what good is, revive some of the long-standing criticisms about development and humanitarianism as preserving Eurocentric systems of knowledge and, ultimately, the coloniality of power.

These observations are accentuated by the geographic imbalance of AI4SG projects. A comprehensive survey of over 1000 AI4SG applications between 2008 and 2020 found that the overwhelming majority of projects were US-based, while the few projects that focused on Africa were run by US researchers (Shi et al., 2020, p. 45). While there are AI4SG projects, including chatbots, which are designed to engage community members (Luria et al., 2020), the majority of Global South projects appear to be top-down (Shi et al., 2020). Computation, and the classifications on which it is based, mirror the values of its coders (Gebru, 2020) which is one of the reasons why processes of automation reproduce and heighten inequality (Eubanks, 2017) and racial and gendered discrimination (Benjamin, 2019; Noble, 2018).

A decolonial approach to AI for social good, doesn't just reveal the power asymmetries underpinning development and humanitarian projects. It also highlights the fact that AI is part of larger genealogies of enumeration, measurement and classification that were originally developed by imperial powers to control colonial subjects (Appadurai, 1993). Biometrics, which in its contemporary iteration is based on AI, was first developed in India as part of the British Empire's efforts to control colonial subjects (Sengoopta, 2003). The colonial genealogies of science are also evident in the way former colonies are used as laboratories for experimentation (Petryna, 2009), a legacy echoed in some AI4SG pilots. Cloaking such practices in the moral discourse of the 'social good' echoes Mbembe's observation that morality was historically used to rationalise colonial practices (2017, p. 11).

A decolonial approach to 'AI for good' shifts the emphasis from questions of ethics to questions of power. The overarching question this article asks is whether humanitarian chatbots, as exemplars of AI4SG, can contribute to challenging inequalities in the global context. This divides into further sub-questions. Do chatbots in humanitarian settings fulfil the promise of intelligence in their ability to hold a conversation? Do chatbots contribute to improved accountability and communication with affected communities? Are there any harms relating to chatbot communication in humanitarian settings? Who benefits from such initiatives and do chatbots revive the colonial legacies on which both humanitarianism and technology are based?

Studying chatbots as part of humanitarian innovation

The article is part of a larger study on digital innovation, data and automation in the humanitarian sector. This is an ethnographic, mixed methods study including interviews, participant observation in spaces of innovation and digital ethnography. In total, I

conducted 45 interviews with seven groups of stakeholders (including humanitarian officers, donors, entrepreneurs and business representatives, digital developers, government representatives, consultants, volunteers).⁷ This was a purposive sample: interviewees were selected in terms of their expertise and its relevance to the research questions. Effort was made to speak to people at different levels within the various organisational hierarchies. Although not all 45 interviews addressed chatbots directly, the present article incorporates insights from the analysis of all interviews regarding digital innovation in humanitarian operations. The research was informed by the ESRC's Framework for Research Ethics.⁸ Interviews were conducted between July 2016 and June 2020 and took place online (via video platforms), or in person (in London, Cambridge, UK; Athens, New York and Washington, DC). Interviews lasted 60–90 min on average and were recorded, transcribed, coded and analysed thematically. All interviews were anonymised and are presented here in an aggregate form in order to protect the anonymity of expert interviewees. For example, when I refer to aid sector participants this may include humanitarian workers or consultants. The article also draws on fieldnotes I took during participant observation at 12 industry events and hackathons.

The article also draws on digital ethnography (Pink et al., 2016): I have consulted blogs, websites, podcasts, videos of talks and relevant events, policy documents, social media content as well as a close reading of documentation regarding the development of chatbots.⁹ Finally, I draw on my own interactions with open access humanitarian chatbots. This autoethnographic exploration, common in digital ethnographies (Barassi, 2020) allowed me to experience first-hand the platform interface.

Humanitarian chatbots: an overview

Chatbots have been developed by humanitarian organisations such as UNHCR and the WFP as well as by private companies such as X2AI, which launched the psychotherapy chatbot 'Karim', or 'Refugee Text', which launched an eponymous chatbot in 2016 to respond to the informational needs of refugees arriving into Europe.¹⁰ Although private initiatives such as 'Refugee Text' received considerable media attention and recognition after being included in a 2017 exhibition at the Design Museum in London, the take up appears to be very minimal.¹¹ Humanitarian agencies have explored the potential of chatbots for information dissemination, communication with communities and accountability to affected people. Another important motivation for humanitarian agencies is cost cutting and efficiencies. For example, the UNHCR Innovation Service, in collaboration with Facebook, developed a chatbot for the UNHCR office in Jordan in order to help with the high volume of calls and the associated costs. The initial pilot (2017–2018) uncovered many issues, including a number of legal and data protection issues, which led UNHCR to decide to withdraw the bot.¹²

The World Food Programme (WFP), through its Mobile Vulnerability and Analysis Unit (mVAM), has been at the frontline of chatbot development in the humanitarian sector since at least 2016. The first chatbot pilot aimed to reach displaced communities and conduct a mobile food security survey. That chatbot ran on the messaging platform Telegram, but subsequent versions used Facebook because it was an already popular platform among the target communities.¹³ These earlier chatbots were developed with the Artificial Intelligence Development Assistance (AIDA) builder which allows the uploading of

questionnaires and spreadsheets and streamlines the collection of structured data. The WFP mVAM unit states that the benefit of disseminating a survey via a chatbot isn't only the 'low cost', but that the method can 'provide a great incentive for people to complete [the] surveys and [provide] feedback on the services we provide'.¹⁴ The mVAM unit identified three core uses for chatbots: information dissemination, simple data collection, and feedback and complaints systems.¹⁵

Later versions of the WFP chatbot were developed using CHITCHAT which uses Natural Language Processing capabilities (NLP) to produce more 'natural' conversations. The CHITCHAT chatbot was deployed through lightweight websites as opposed to Facebook and was tested in the Kakuma and Kalobeyei refugee camps in Western Kenya.¹⁶ According to a review case study, the pilot was hampered by serious access issues in the two camps, which are indicative of digital inequality patterns among displaced people.¹⁷ In Kakuma 90% of households had access to either a basic mobile phone or smartphone and internet was available in some locations. By contrast, in Kalobeyei fewer than 20% of households had a phone while charging was difficult and expensive. While many young people claimed to have Facebook accounts, few could access them as they had not yet obtained a Kenyan SIM card and struggled to do so without the proper identification (DIAL, 2018, p. 12). In both camps, mobile internet data were often unaffordable with refugees struggling to pay expensive rates without any sources of income. In other pilots, the WFP relied on the Facebook programme 'Free Basics' in order to navigate access issues.¹⁸ The latest iteration of the WFP chatbot is 'Agrochatea', an automated chatbot that provides wholesale market prices for agricultural produce using Natural Language Processing capabilities. Agrochatea was developed by the WFP, the Peruvian government and the Centre for Innovation at Leiden University and is also accessible through a dedicated lightweight website that can run on both smartphones and basic internet-enabled phones.¹⁹

The analysis that follows refers to the above examples which are representative of similar developments in the humanitarian sector.

Artificial conversations, unintelligent bots

On close inspection, several humanitarian chatbots confirm the conclusion of a recent report that they are 'simple functional bots that stick reliably to a predefined script' (ICRC et al., 2017). Even CHITCHAT which uses Natural Language Processing offers answers from a range of pre-defined options. According to the WFP website, the CHITCHAT bot allows for 'more natural conversations' ... 'as users chat with the bot using their own questions rather than keywords and the bot answers with information relating to their query'.²⁰ My interaction with the bot,²¹ on 10 separate occasions between February and December 2020 exposed several limitations. Every time I asked a complex question such as 'why do people in Kakuma and Kalobeyei receive different cash assistance' the chatbot responded 'Sorry I don't understand that' and took me back to the first step. Below is a transcript of my longest conversation with the chatbot recorded on 2 December 2020. *Bamba Chakula* means 'get your food' in Kiswahili and in this context refers to the WFP's cash distribution programme in the refugee camps in Kenya.

Chatbot: Hi There!

This is an automated chatbot from the World Food Programme. I can help you with questions about WFP programmes, this chatbot project, Bamba Chakula and General Food Distributions.

You can ask a question or choose from the topics below. How can I help?

- General Food Distributions
- WFP Programmes
- Bamba Chakula
- Chatbot Project

Researcher: Bamba Chakula

Chatbot: The Next Bamba Chakula will arrive on 27 May [emphasis added].

Is there anything in particular that you would like to know?

You can ask a question or choose from the topics below.

- Info about Bamba
- Get Balance
- Amount

Researcher: how much will I receive?

Chatbot: First are you in Kalobeyei or Kakuma?

Researcher: Kakuma

Chatbot: It also depends on the size of your family.

How MANY PERSONS are there in your household?

Researcher: 1

Chatbot: In Kakuma, Bamba Chakula is 400 KES per person.

Researcher: Why do people in Kakuma and Kalobeyei receive different amounts?

Chatbot: Sorry, I don't understand that.

This is chatbot that can help you with questions about Bamba Chakula, General Food Distributions (GFD), and information about WFP programmes in Kakuma and Kalobeyei.

How can I help?

- GFD
- Bamba Chakula
- WFP programmes

While in this example CHITCHAT was able to recognise phrases and provide useful answers, it wasn't able to decode my more complex question. Recall that this is the longest conversation I held with the chatbot – most interactions ended after my first or

second question. On this occasion, the chatbot also provided me with wrong information regarding the date. Given *Bamba Chakula* is a monthly distribution, the earliest date after December cannot be May.²² According to one of the aid sector interviewees: ‘these chatbots aren’t very different from the surveys that already go out’ via traditional means [...] what’s different is speed, efficiency’. Given that humanitarian contexts give rise to complex problems that do not fit predetermined categories, the limited degree to which a chatbot can respond is a problem.

This finding shouldn’t be surprising. Humanitarian chatbots follow the established tradition of humanitarian feedback policies. By feedback most humanitarian organisations refer to comments about specific interventions that they have initiated. This narrow definition of feedback became apparent in an earlier ethnography which examined the humanitarian response to Typhoon Haiyan in the Philippines.²³ In that context, aid agencies solicited feedback in relation to the delivery of particular services. For example, an agency which had rolled out a sanitation initiative, would only collect feedback on that particular project – not on other issues which may have been more pressing among affected people. As a result, more global comments regarding people’s experiences were not systematically recorded (Madianou et al., 2016). In the rare occasions when more general issues were recorded, they were not responded to as they didn’t fit into the remit of each particular agency or cluster which created a disconnect between affected people and humanitarian organisations (Madianou et al., 2016). An aid worker interviewee described this problem:

Humanitarian organisations are constantly sent information that doesn’t fit neatly into buckets. So people will say, I got information about political violence, but we’re supposed to be a neutral humanitarian project, what do you do with that? So there’s all these kinds of situations where it’s not going to fit neatly.

Chatbots represent the logical step in the tradition of digitised feedback in the humanitarian sector. Rather than facilitating a meaningful engagement with communities, feedback is instrumentalised and reduced to formulaic interactions and templates. Chatbots also prioritise information dissemination, a process which doesn’t require the complex infrastructure of AI. As an NGO participant remarked:

All chatbots are about pushing information out. Even ‘Refugee Text’ is: ‘tell us your status and we’ll give you some information on that basis’. Maybe at best it’s tailored information, but it’s not a conversation. [...] Participation is hard to do. It’s easy to push out information.

These observations are compounded by another limitation of the chat format which privileges short text rather than longer questions or explanations. In my earlier research with communities affected by Typhoon Haiyan, where feedback was collected through SMS texting, the character limit of that particular channel prevented participants from expressing complex concerns for example, regarding their exclusion from aid distribution lists (Madianou et al., 2016). While the CHITCHAT bot has no character limits, the genre of Instant Messaging which it emulates favours brief inputs. The brevity of text interactions is also necessary as the chatbot is more likely to respond to very short commands rather than complex sentences. According to one of the aid sector interviewees:

there's going to be a lot of messy, complicated situations all where things change very fast where it's just not appropriate to give people automated answers and feel it's accurate. Automation in this approach has the potential to be very dehumanising.

Automation is a technology of distancing. As the term 'nonhuman' implies, automation streamlines processes at a distance, thus separating a person from the consequences of their actions. Automation can also be a frustrating experience when a chatbot doesn't understand one's questions and abruptly restarts the conversation. Anyone who has struggled with automated responses in customer service departments, when no predetermined answer seems to address one's query, can probably sympathise with these observations.

Although it's not the aim of this article to provide a systematic evaluation of specific apps, their narrow topic remit and their limited capacity to respond to questions that deviate from the standard script raises concerns about their usefulness. Another issue is whether the topics covered resonate with the actual needs of communities. The example of 'Agrochatea', the WFP chatbot that provides real time information of market produce, evokes the study by Burrell and Oreglia (2015) on 'the myth of market-price information'. Drawing on two ethnographies in China and Uganda, Burrell and Oreglia contested the oft-repeated notion that farmers use mobile phones to access market price information. While their ethnographies revealed several creative uses of mobile phones, there was no evidence that people used their phones to find out about market prices and therefore increase the efficiency of their transactions. The authors question the myth of 'market price information' which has become a 'shorthand for the potential of digital technologies to empower low-income communities in the global south' (Burrell & Oreglia, 2015, p. 271). Their study illustrates how predefined notions of 'social good' in development projects do not correspond with the experiences of local communities. One wonders whether 'Agrochatea' is the product of the 'market information mantra', or whether the chatbot's developers responded to genuine community demand for this kind of information.

When 'AI for good' can be harmful

While there are clear questions regarding the intelligence and usefulness of chatbots, this doesn't mean there are no potential harms. The first issue to consider is the question of data safeguards especially given the heightened vulnerability of humanitarian settings. The fact that many of the above chatbots use existing platforms such as Facebook Messenger or Whatsapp means that the safeguarding and privacy of the data of humanitarian subjects is outsourced to Silicon Valley companies. By outsourcing services to Facebook, humanitarian agencies inherit the concerns surrounding the business model of social media companies whereby users' data are extracted for profit (Zuboff, 2019). By relying on Facebook or other messaging apps for essential services, humanitarian agencies extend their remit ('the imperative to do good' and 'do no harm') to private technology companies (McDonald, 2019). In the case of the AIDA chatbot, the use of Facebook Messenger wasn't the result of a formal partnership which meant that the WFP had no control over the safeguarding of the data and metadata, which are equally sensitive when dealing with persons of concern (ICRC and Privacy International, 2018). The same

applies to the case of the psychotherapy chatbot, Karim. Even in cases when humanitarian organisations sign formal partnerships with big technology companies it is not clear ‘what leverage they have to ensure the enforcement of that control’ (McDonald, 2019, p. 4). This asymmetry is even more pronounced in the aforementioned chatbot cases where no contract with Facebook has been signed.

Although not explicitly stated, such privacy and data safeguarding concerns may have contributed to the shifting of the WFP chatbots to a dedicated lightweight website. Issues of safeguarding, of course, remain relevant even with the bespoke light website. Humanitarian organisations are often the target of cyber-attacks.²⁴ The vulnerability from potential data breaches is amplified by the aid agencies’ ‘tendency to maximize data collection just in case it’s needed in the future’ a point that came up in several interviews. Ironically, if the lightweight website evades the questionable data extraction practices of social media platforms, it may be less accessible to local communities. The reason, after all, for using Facebook Messenger for the original WFP chatbot, was because it was already part of people’s communication environments. An aid sector interviewee expressed their reservation about the success of bespoke apps or websites as ‘they are not part of [the refugees’] communication ecologies’. Facebook is also implicated in access issues as the WFP encourages access to its lightweight websites through Free Basics, a programme criticised for its extractive logics (Nothias, 2020).

Of greater concern is the potential risk of misinformation if, for example, an app is out of date and offers mistaken information as evidenced in my interaction with the CHIT-CHAT bot discussed above. If a refugee is given wrong dates or other out-of-date information the consequences can be severe. As one of the NGO interviewees put it, ‘if a chatbot directs refugees to the wrong meeting point, and it has taken them hours and money to get there this can be very problematic’. Misinformation, however unintentional, can be at odds with the humanitarian imperative of ‘do no harm’. This raises the issue of responsibility in ‘human-machine communication’ (Gunkel, 2018). Who is responsible if a chatbot disseminates out-of-date information? Who can be held accountable? This is why the same interviewee remarked: ‘You’ve got to have a human there, who can step in. But given these apps are driven by efficiency concerns, I’m not sure they will be prepared to resource it’.

The lack of meaningful communication reveals that, just like with several other humanitarian innovations, the purpose of chatbots is to maximise efficiencies and savings which is openly acknowledged in the blogs and websites which accompany the UNHCR and WFP chatbots.²⁵ Aiming for efficiency isn’t necessarily harmful, unless the logic of efficiency and cost-cutting reduces interactions to their barest instrumental forms. This can be problematic if affected people are promised ‘participation’ and ‘accountability’ as part of AAP initiatives, but are ultimately only offered a box-ticking exercise. In the Typhoon Haiyan research discussed earlier, the formulaic nature of feedback mechanisms was felt deeply among affected people, many of whom felt disconnected from the aid agencies especially when their messages were not responded to (Madianou et al., 2016). If chatbots represent a similarly monological proposition, they will only provide an illusion of accountability.

Such concerns are likely to be exacerbated if the logic of efficiency is accompanied by the extraction of data for audit purposes. As discussed above, the marketisation of the humanitarian field has created pressures for constant reporting in order to satisfy the

donors' demand for evidence of impact and audit trails which are then used to justify the renewal of funding. One of my aid sector interviewees described chatbots and feedback apps as 'sucking apps ... because they suck from communities'. This illustrates that humanitarian subjects pay for aid with their own data which are extracted to legitimate humanitarian projects. Extraction also occurs through the value accrued by the experimentation with untested technologies in humanitarian settings – something I shall explore in the following section.

Reproducing coloniality

Chatbots, like all forms of algorithmic culture, contribute to defining the parameters of humanitarian aid. The algorithms of natural language processing 'condense plural possible pathways to a single output' (Amoore, 2020, p. 17) thus foreclosing other possibilities. When considering what kind of knowledge, the chatbots produce, it is important to remember what an aid sector interviewee noted: chatbots – just like most innovations – 'have western bias in their design. [They] are tested on people in their vicinity'. Even though the WFP foodbot was piloted in Kenya, it was conceived of, and designed in, hackathons and labs far removed from the reality of the refugee camps. For example, the AIDA foodbot was developed in a hackathon organised by the global marketing and data analytics company Nielsen on behalf of mVAM in New York City in January 2017. The hackathon was attended by 'developers, students, volunteer hackers and Nielsen staff' (DIAL, 2018, p. 9) – presumably all residents of New York or other global cities. This reflects a point that many of my interviewees agreed on: hackathons are mostly white and male with no refugee participation. This observation confirms my own experience of several hackathons where no refugees or members of affected communities took part. A survey of 49 AI applications in the humanitarian sector confirms that all are led by organisations in the Global North (predominantly North America and Europe).²⁶

Given the design and management of AI4SG projects is located in the Global North it is not surprising that the values which inform some of these platforms are western-centric. For example, the notion of accountability that informs policies of 'Accountability to Affected People' is a Eurocentric notion that often has little purchase in local contexts. Earlier research in the Philippines, reported that humanitarian aid was filtered through local norms of debt of gratitude, which itself was grounded in colonial legacies and local patronage politics (Madianou et al., 2016). The local idioms around humanitarian aid as gratitude were at odds with the assumptions of humanitarian accountability policies which usually translate narrowly into feedback collection. In the Philippine study, local norms prevented members of affected communities from any public expressions of criticism even when their private accounts were highly critical of aid operations.

The language in which chatbots operate is vital here. While some of the WFP chatbots (e.g., Agrochatea) are translated into local languages, others are not (e.g., CHITCHAT bot in Kakuma pilot). Defending that decision, the mVAM team argued that most of their target users already spoke English.²⁷ However, communication in English excludes older users who were also experiencing additional access issues. Translating the bot into the languages (Dinka, Kiswahili or Somali) spoken in the Western Kenya refugee camps is a vital step towards inclusivity. Decolonial writers like Wa Thiong'o have stressed how 'language carries culture and culture carries [...] the entire body of values through which

we perceive ourselves and our place in the world' (2004, p. 16). Language serves as tool that sustains global hierarchies.

The importance of cultural sensitivity is even more vital for chatbots like the psychotherapy one ('Karim') which opened the article. Awareness of cultural norms and idioms is crucial in psychotherapy especially as sociocultural factors contribute to mental health and illness. Given the cultural specificity of emotions, a deep understanding of cultural codes and norms is necessary for psychotherapy to be successful. The same applies to gender, sexuality and social class which, too, are culturally specific. Cultural sensitivity is not just achieved by translating script into different languages. What matters is understanding how emotions like shame manifest differently depending on the cultural context which requires a deep immersion in a local culture. Cultural sensitivity can also be achieved by the inclusion of local trained psychotherapists who can intervene when necessary. 'Karim' is a pared down version of 'Tess', X2AI's behavioural coaching chatbot that addresses depression and anxiety through existing instant messaging apps (Stephens et al., 2019). While 'Tess' serves as a therapeutic tool that supplements, rather than replaces the role of a therapist,²⁸ the roll out of 'Karim' was not accompanied by the parallel availability of professional psychotherapists.

Such asymmetries in the roll out of similar platforms in the US and Lebanon echo arguments about scientific experimentation in former colonies (Petryna, 2009), which I discussed earlier in the article as part of colonial genealogies. Why are trained psychotherapists necessary to support US users of chatbots, but not refugees in Lebanon? Why are refugees in Lebanon receiving a 'pared down' version of the chatbot, when the mental health issues they face can be more acute due to the trauma of war? Private companies involved in AI4SG projects often extract value from experimentation with untested technologies. By piloting technologies such as chatbots, private companies extract value to improve the design of their products and crucially, to generate hype around their innovations, which translates into attention and ultimately profit. As one aid sector participant put it:

We have this interesting situation that we're in [...] when companies invest a lot of time and effort for products that they haven't yet found a market for. [...] And there's a lot of effort trying to create hype. The humanitarian sector is one of the ways of creating headlines. This is essentially advertising for companies.

The psychotherapy chatbot aimed at Syrian refugees in Lebanon is a good example of how a particular pilot can generate interest in a company. X2AI received significant publicity for its non-profit work with feature articles in publications such as the *New Yorker* and *The Guardian* among others²⁹, which it might not have received for its other products. Experimentation in vulnerable settings can translate into publicity and ultimately, profit. The hype around chatbots, just like the hype around other innovations such as blockchain (Madianou, 2021) is one of the driving forces in the development and roll out of humanitarian innovations especially those involving the private sector. This is a point where the logics of technological solutionism and capitalism meet the logic of humanitarian marketisation creating a dynamic described as technocolonialism (Madianou, 2019a). The desire for efficiencies and audit trails, often demanded by donors, finds the perfect match in the logic of solutionism which is often pushed by large technology companies that seek branding opportunities and visibility for their products.

The fact that chatbots are designed in the Global North and implemented in the Global South, as is typical for AI4SG humanitarian projects; the fact that projects do not always have the linguistic, or cultural sensitivity required; and finally, the extraction of value from experimentation with untested technologies all contribute to the reworking of the colonial legacies of humanitarianism and AI4SG.

The following section puts forward an additional way in which humanitarian AI4SG projects reproduce power asymmetries.

A technology of enchantment?

The power of ‘AI for good’ is most evident in the ways in which it produces a sense of enchantment. I here draw on Gell’s seminal work on ‘The Technology of Enchantment and the Enchantment of Technology’ (1992). Gell was interested in how artifacts ‘cast a spell’ over people by ‘functioning as weapons in psychological warfare’ (1992, p. 44). The example he uses is the intricate and imposing canoe prow board from the Trobriand islands which is exquisitely crafted to dazzle the Trobrianders’ overseas partners so that they will ‘take leave of their senses and offer more valuable shells or necklaces to the members of the expedition than they would otherwise be inclined to do’ (Gell, 1992, p. 44). The prow board, with its intricate engraving and adornments, enchants and helps the Trobrianders achieve their goal to maximise the value of their trade. The canoe-board ultimately confers magical prowess on the owners of the canoe whose power and status are, in turn, heightened.

The magical efficacy of technology is achieved by hiding the work through which it was produced (Gell, 1992). In other words, ‘the enchantment of technology’ (the hold technologies have on us) depends on ‘the technology of enchantment’ – the actual making of artifacts (Gell, 1992). This crafting typically includes erasure, or in the case of communication technologies, ‘immediacy’ (Bolter & Grusin, 2000). Immediacy refers to the capacity of mediation to erase its own work, to appear natural, or unmediated. The opacity of algorithms and systems of automation is a perfect example of ‘immediacy’, which is one of the reasons why Chun (2011, p. 87) describes software as ‘magic’. If chatbots, however limited in their capabilities, are examples of the ‘enchantment of technology’ (Gell, 1992) then that makes them potentially more powerful. Their power relates to the way they are fetishised as powerful, or enchanted; not necessarily in their efficaciousness. Ultimately, the enchanted object, in this case the chatbot, confers power to its owner: the humanitarian organisations or the private companies and their philanthropic departments. When the ‘Refugee Text’ chatbot was selected for inclusion in the 2018 Beazley Designs of the Year exhibition at the Design Museum in London, it represented a clear case of the ‘enchantment of technology’. Through its selection and display in a high-profile exhibition, the chatbot acquired enchanted status as an item of cultural and social significance. Similarly, the extensive publicity that the psychotherapy chatbot received, conferred visibility to its company. The ‘enchantment of technology’ also explains why ‘AI for good’ is increasingly appropriated as a marketing and branding strategy.

Placing an ‘enchanted technology’ in a very asymmetrical context can potentially amplify existing power imbalances. By conferring authority to humanitarian organisations and for-profit companies, chatbots rework and revitalise the existing asymmetries

of humanitarianism and reproduce the coloniality of power. Humanitarian AI may not be an example of ‘psychological warfare’ (Gell, 1992), but it certainly has the potential to create hierarchies and boundaries between the owners and recipients of technology. In the case of the Design Museum exhibition, the clear winners are the chatbot developers who gained visibility and distinction through their inclusion in a high-profile exhibition. This contrasts with the invisibility of refugees, who remain stuck in camps in Greece and Italy and whose plight provided the raw material for the elevation of the chatbot designers.

Conclusion

The article analysed chatbots as emblematic of ‘AI for social good’ initiatives in the humanitarian field. The article develops an interdisciplinary framework that brings together the decolonial critique of humanitarianism and critical algorithm studies and draws upon a qualitative study of humanitarian innovation in order to unravel some of the assumptions behind the ‘AI for social good’ phenomenon. This approach shifts the emphasis from questions of ethics to questions of power. Recognising that humanitarianism, as well as AI, are steeped in unequal relations of power, the analysis focused on the extent to which chatbots reproduce inequalities and colonial legacies.

The article observed that humanitarian chatbots remain limited in terms of being able to ‘hold a conversation’ which is how chatbot ‘intelligence’ is often defined. Even if they are not as powerful as is often suggested in popular discourses, chatbots still have powerful consequences. The article identified significant risks regarding data safeguards and misinformation which can have deleterious implications for refugees and other vulnerable people. Chatbots reduce participation and feedback to its barest instrumental forms which creates disconnects between affected communities and aid agencies. This disconnect is compounded by the extraction of value from data for the purposes of audit and humanitarian metrics. The distancing of automation, which is often referred to as ‘nonhuman computation’, has the potential to dehumanise interactions in the humanitarian context whilst claiming to be objective and scientific.

Because humanitarian chatbots are designed in the Global North and implemented in the South, and because of the concerns regarding linguistic or cultural sensitivity, I have argued that such innovations rework the colonial legacies of humanitarianism and AI4SG. The extraction of value from the experimentation with untested technologies further reflects colonial genealogies. The reworking of the power relations of coloniality and humanitarianism is most evident in the ‘enchantment of technology’ (Gell, 1992). By ‘casting a spell’, technologies amplify social hierarchies and confer status on the technology designers. In the case of AI, the enchantment of technology is achieved through the occlusion of the work of algorithms and the parallel obscuring of the underlying relations of inequality. The notion of the ‘enchantment of technology’ (Gell, 1992) suggests that AI4SG, together with other sociotechnical assemblages, may play a constitutive role in the reworking of colonial legacies.

Not all chatbots are the same. For example, there are differences between public and private innovations. Yet on the whole, the structural issues regarding power asymmetries, knowledge production and enchantment apply to all the examples discussed here. The distinction between good or bad forms of AI misses the point. All forms of AI contribute

to knowledge production and, ultimately, the way the world is perceived. As Amoore (2020) argues, there are no good, or bad algorithms; all algorithms define the parameters of social life. Similarly, it is not meaningful to define AI as good or bad. But because all forms of AI perform a worldmaking role, all forms of AI can potentially be harmful.

Notes

1. 'The chatbot will see you now': <https://www.newyorker.com/tech/annals-of-technology/the-chatbot-will-see-you-now>
2. For a list of all UN sustainable development goals see: <https://www.un.org/sustainabledevelopment/sustainable-development-goals/>
3. Biometric technologies depend on artificial intelligence in the capturing, segmenting and indexing of biometric data, as well as in the matching of biometric records against large databases. Because of the way they involve states and policies of securitisation, biometrics are discussed in a separate publication (Madianou, 2019b, Madianou, 2021).
4. <https://aiforgood.itu.int/>
5. https://reliefweb.int/sites/reliefweb.int/files/resources/GHA%20report%202019_0.pdf; <https://www.unhcr.org/globaltrends2019/>
6. see notes 2 and 4 for the UN SDGs and AI for Good Summit.
7. These distinctions aren't always clear cut, partly because participants may have, or have had multiple roles, for example as aid workers and consultants; or entrepreneurs and digital developers.
8. <https://escr.ukri.org/funding/guidance-for-applicants/research-ethics/>
9. Hundreds of sources were collected as part of the wider project which have shaped my understanding of humanitarian innovation. The article references 28 of those sources, details of which can be found in the notes.
10. <http://www.refugeetext.org/>
11. Beazley Designs of the Year, The Design Museum: <https://designmuseum.org/exhibitions/beazley-designs-of-the-year/digital-20x/refugee-text> 'How Chatbots are helping refugees navigate a new life', The Guardian: <https://www.theguardian.com/public-leaders-network/2016/oct/17/refugee-information-facebook-social-media-tech-help> Last accessed 26 November 2020.
12. UNHCR Innovation Service: Chatbots in humanitarian settings: revolutionary, a fad or something in-between? <https://www.unhcr.org/innovation/chatbots-in-humanitarian-settings-revolutionary-a-fad-or-something-inbetween/> Last accessed 26 November 2020.
13. WFP, mVAM: Introducing our chatbot: <http://mvam.org/2016/07/21/introducing-our-chatbot/>
14. WFP, mVAM: Introducing our chatbot: <http://mvam.org/2016/07/21/introducing-our-chatbot/>
15. <http://mvam.org/2019/05/29/hi-chatbot-what-have-you-been-up-to-lately/>
16. The Kakuma Refugee Camp (established in 1992) in the Turkana district of Western Kenya is one of the world's largest refugee camps. Refugees are mostly from South Sudan and Somalia. The neighbouring settlement of Kalobeyei was established in 2015 for more recent arrivals. By July 2020 Kakuma and Kalobeyei had a population of 196,666 refugees. <https://www.unhcr.org/ke/kakuma-refugee-camp>
17. <http://mvam.org/2017/09/06/how-can-chatbots-help-us-respond-to-humanitarian-crisis/>
18. https://soundcloud.com/wfp_vam/vam-talks-episode-4/sets
19. WFP, mVAM: ¡Hola Agrochatea! <http://mvam.org/2019/08/27/hola-agrochatea/>
20. Hi Chatbot! Where have you been to lately? <http://mvam.org/2019/05/29/hi-chatbot-what-have-you-been-up-to-lately/> Last accessed 16 December 2020.
21. The chatbot can accessed here: <https://speakto.ita.wfp.org/>
22. Bamba Chakula: a saviour in times of reduced rations in Kenya. <https://medium.com/world-food-programme-insight/bamba-chakula-a-saviour-in-times-of-reduced-rations-fd6604b4df44> Last accessed 16 December 2020.

23. Typhoon Haiyan made landfall on 8 November 2013.
24. Among the several reports of data breaches and cyberattacks involving humanitarian organisations, the most significant occurred in August 2019 involving multiple UN agencies. <https://www.thenewhumanitarian.org/investigation/2020/01/29/united-nations-cyber-attack> Last accessed 26 November 2020.
25. UNHCR Innovation Service: Chatbots in humanitarian settings: revolutionary, a fad or something in-between? <https://www.unhcr.org/innovation/chatbots-in-humanitarian-settings-revolutionary-a-fad-or-something-inbetween/> Last accessed 26 November 2020.
26. <https://centre.humdata.org/catalogue-for-predictive-models-in-the-humanitarian-sector/>
27. <http://mvam.org/2017/09/06/how-can-chatbots-help-us-respond-to-humanitarian-crisis/>
28. <https://www.x2ai.com/>
29. Articles in the *Guardian* <https://www.theguardian.com/technology/2016/mar/22/karim-the-ai-delivers-psychological-support-to-syrian-refugees>, *The New Yorker* <https://www.newyorker.com/tech/annals-of-technology/the-chatbot-will-see-you-now>, Reuters <https://www.reuters.com/video/watch/id117741381>, *Business Insider* <https://www.businessinsider.com/psychotherapy-bot-in-middle-east-2016-3?r=US&IR=T> among others.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Mirca Madianou (PhD London School of Economics and Political Science) is Reader in the Department of Media, Communications and Cultural Studies at Goldsmiths, University of London. Her current research focuses on the social consequences of communication technologies, data and automation in the global south especially in relation to migration and humanitarian emergencies. She is the author of *Mediating the Nation: news, audiences and the politics of identity*, and *Migration and New Media: transnational families and polymedia* (with Daniel Miller). [email: m.madianou@gold.ac.uk].

References

- Amoore, L. (2020). *Cloud ethics. Algorithms and the attributes of ourselves and others*. Duke University Press.
- Anwar, M. A., & Graham, M. (2020). Digital labour at economic margins: African workers and the global information economy. *Review of African Political Economy*, <https://doi.org/10.1080/03056244.2020.1728243>
- Appadurai, A. (1993). Number in the postcolonial imagination. In C. Breckenridge & P. Van der Veer (Eds.), *Orientalism and the postcolonial predicament* (pp. 314–339). University of Pennsylvania Press.
- Bakardjieva, M. (2015). Rationalizing sociality: An unfinished script for socialbots. *The Information Society*, 31(3), 244–256. <https://doi.org/10.1080/01972243.2015.1020197>
- Barassi, V. (2020). *Child, data, citizen*. MIT Press.
- Benjamin, R. (2019). *Race after technology*. Polity.
- Bolter, J. D., & Grusin, R. (2000). *Remediation*. MIT Press.
- boyd, d., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Broussard, M. (2018). *Artificial unintelligence*. MIT Press.
- Burrell, J., & Oreglia, E. (2015). The myth of market price information: Mobile phones and the application of economic knowledge in ICTD. *Economy and Society*, 44(2), 271–292. <https://doi.org/10.1080/03085147.2015.1013742>

- Calhoun, C. (2008). The imperative to reduce suffering. In M. Barnett & T. G. Weiss (Eds.), *Humanitarianism in question* (pp. 73–97). Cornell University Press.
- Charlson, F., van Ommeren, M., Flaxman, A., Cornett, J., Whiteford, H., & Saxena, S. (2019). New WHO prevalence estimates of mental disorders in conflict settings: A systematic review and meta-analysis. *The Lancet*, 394(10194), 240–248. [https://doi.org/10.1016/S0140-6736\(19\)30934-1](https://doi.org/10.1016/S0140-6736(19)30934-1)
- Chouliaraki, L. (2013). *The Ironic Spectator: Solidarity in the age of post-humanitarianism*. Cambridge, UK: Polity.
- Chun, W. H. K. (2011). *Programmed visions: Software and memory*. MIT Press.
- Craig, C., Cave, S., Dihal, K., Dillon, S., Montgomery, J., Singler, B., & Taylor, L. (2018). *Portrayals and perceptions of AI and why they matter*. The Royal Society. Retrieved December 15, 2020.
- Crawford, K., & Joler, V. (2018). Anatomy of an AI System: the Amazon Echo as an anatomical map of labour, data and planetary resources. Retrieved December 16, 2020, from <https://anatomyof.ai/>
- Digital Impact Alliance (DIAL). (2018). *Foodbot and the AIDA Chatbot builder: case study*. <https://digitalimpactalliance.org/wp-content/uploads/2019/03/mVAM.pdf>
- Eubanks, V. (2017). *Automating inequality. How high-tech tools profile, police and punish the poor*. St Martin's Press.
- Floridi, L., Cows, J., King, T. C., & Taddeo, M. (2020). How to design AI for social good: Seven essential factors. *Science and Engineering Ethics*, 26(3), 1771–1796. <https://doi.org/10.1007/s11948-020-00213-5>
- Gebru, T. (2020). Race and gender. In M. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford handbook of ethics of AI* (pp. 253–270). Oxford University Press.
- Gehl, R. W., & Bakardjieva, M. (2017). Socialbots and their friends. In R. W. Gehl & M. Bakardjieva (Eds.), *Socialbots and their friends: Digital media and the automation of sociality* (pp. 1–16). Routledge.
- Gell, A. (1992). The technology of enchantment and the enchantment of technology. In J. Coote & A. Shelton (Eds.), *Anthropology, art and aesthetics* (pp. 40–66). Clarendon.
- Gunkel, D. (2018). Ars ex machina: Rethinking responsibility in the age of creative machines. In A. Guzman (Ed.), *Human-Machine communication* (pp. 221–236). Peter Lang.
- International Committee of the Red Cross (ICRC) and Privacy International. (2018). The humanitarian metadata problem: 'Doing no harm' in the digital era.
- International Committee of the Red Cross, The Engine Room and Block Party. (2017). *Humanitarian futures for messaging apps*. ICRC.
- Kember, S., & Zylinska, J. (2012). *Life after new media. Mediation as a vital process*. MIT Press.
- Krause, M. (2014). *The good project*. Chicago University Press.
- Lester, A., & Dussart, F. (2014). *Colonization and the origins of humanitarian governance*. Cambridge University Press.
- Luria, M., Seering, J., Forlizzi, J., & Zimmerman, J. (2020). Designing chatbots as community-owned agents. In *Proceedings of the 2nd Conference on Conversational User Interfaces (CUI '20)* (pp. 1–3). Article 14. <https://doi.org/10.1145/3405755.3406134>
- Madianou, M. (2019a). Technocolonialism: Digital innovation and data practices in the humanitarian response to refugee crises. *Social Media & Society*, 5(3), 1–13. <https://doi.org/10.1177/2056305119863146>
- Madianou, M. (2019b). The biometric assemblage: Surveillance, experimentation, profit, and the measuring of refugee bodies. *Television & New Media*, 20(6), 581–599. <https://doi.org/10.1177/1527476419857682>
- Madianou, M. (2021). Technological futures as colonial debris: Tech-for-good as technocolonialism. In Zylinska, J. (Ed.), *The Future of Media*. London, UK: Goldsmiths Press.
- Madianou, M., Ong, J. C., Longboan, L., & Cornelio, J. S. (2016). The appearance of accountability: Communication technologies and power asymmetries in humanitarian aid and disaster recovery. *Journal of Communication*, 66(6), 960–981. <https://doi.org/10.1111/jcom.12258>
- Mbembe, A. (2017). *Critique of black reason*. Duke University Press.

- McDonald, S. (2019, August 12). *From space to supply chain: Humanitarian data governance*. <https://ssrn.com/abstract=3436179> or <http://dx.doi.org/10.2139/ssrn.3436179>
- Natale, S. (2019). If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA. *New Media and Society*, 21(3), 712–728. <https://doi.org/10.1177/1461444818804980>
- Noble, S. (2018). *Algorithms of oppression*. New York University Press.
- Nothias, T. (2020). Access granted: Facebook's free basics in Africa. *Media, Culture & Society*, 42(3), 329–348. <https://doi.org/10.1177/0163443719890530>
- Pasquale, F. (2015). *The Black Box society*. Cambridge University Press.
- Petryna, A. (2009). *When experiments travel. Clinical trials and the global search for human subjects*. Princeton University Press.
- Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2016). *Digital ethnography: Principles and practice*. Sage.
- Quijano, A. (2000). Coloniality of power and eurocentrism in latin america. *International Sociology*, 15(2), 215–232. <https://doi.org/10.1177/0268580900015002005>
- Sengoopta, C. (2003). *Imprint of the Raj: How fingerprinting was born in colonial India*. Macmillan.
- Shi, Z. R., Wang, C., & Fang, F. (2020). Artificial intelligence for social good: A survey. *ArXiv [preprint]*. arXiv:2001.01818 [cs.CY].
- Stephens, T. N., Joerin, A., Rauws, M., & Werk, L. (2019). Feasibility of pediatric obesity and pre-diabetes treatment support through tess, the AI behavioral coaching chatbot. *Translational Behavioral Medicine*, 9(3), 440–447. <https://doi.org/10.1093/tbm/ibz043>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind; A Quarterly Review of Psychology and Philosophy*, 49(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Wa Thiong'o, N. (1981/2004). *Decolonising the mind: The politics of language in African literature*. James Currey.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Zuboff, S. (2019). *The age of surveillance capitalism*. Profile.