International Journal of Communication 15(2021), 343–362 1932–8036/20210005

Giving by Taking Away: Big Tech, Data Colonialism, and the Reconfiguration of Social Good

JOÃO CARLOS MAGALHÃES
Alexander von Humboldt Institute for Internet and Society, Germany

NICK COULDRY
London School of Economics and Political Science, UK

Big Tech companies have recently led and financed projects that claim to use datafication for the "social good." This article explores what kind of social good it is that this sort of datafication engenders. Drawing mostly on the analysis of corporate public communications and patent applications, it finds that these initiatives hinge on the reconfiguration of social good as datafied, probabilistic, and profitable. These features, the article argues, are better understood within the framework of data colonialism. Rethinking "doing good" as a facet of data colonialism illuminates the inherent harm to freedom these projects produce and why, to "give," Big Tech must often take away.

Keywords: datafication, social good, Big Tech, data colonialism, political economy

The COVID-19 crisis has created unprecedented opportunities for those with large data processing resources to claim a privileged position to offer social solutions, whether via contact tracing apps (Newton, 2020), AI in managing scarce health resources (Hao, 2020), or AI-driven population tracking (Lewis, Conn, & Pegg, 2020). Yet the idea of using data-driven computational systems for social benefit predates (and will survive) the coronavirus pandemic. Part of the initiatives previously named digital humanitarianism (Meier, 2015) and Big Data for development (Hilbert, 2016) have been recently bundled under the expression *social good* (International Telecommunication Union, 2020). Though no precise numerical estimate exists of such initiatives, they appear to have grown hugely. Consider the number of academic papers on AI for social good (AI4SG), which, according to one count, increased by more than tenfold, from 18 in 2008 to 246 in 2019 (Shi, Wang, & Fang, 2020). AI4SG projects tackle problems as distinct as diagnosing crop diseases and "empowering" refugees (Chui et al., 2018). But they all hinge on datafication: the conversion of ever more aspects of life into digital data for algorithmic mining and semiautonomous decision making (van Dijck, 2014).

This article explores *what kind* of social good it is that datafication may engender. If the production of social good as a reference point for knowledge and policy is a crucial aspect of how social reality is constructed at all times, the articulation of social good within *particular* regimes of knowledge is as important

João Carlos Magalhães: joao.magalhaes@hiig.de
Nick Couldry: n.couldry@lse.ac.uk
Date submitted: 2020-07-07

today as in relation to the expanding philanthropy and social intervention that accompanied the development of statistical knowledge in 19th century Europe.[1] We discuss specific actors in the vast datafication landscape: Facebook, Google, Microsoft, Amazon, and IBM, whose social good initiatives affect various sorts of vulnerable populations. Drawing mostly on the analysis of their public communications and patents fillings, we propose that Big Tech instantiate a specific kind of social good that applies large-scale commercial datafication technologies to problems that are neither commercial nor necessarily datafiable.

The article first summarizes the critical literature on the topic, contextualizing the datafication for the social good within a broader corporatization of social knowledge. After justifying our methodological decisions, it characterizes the sort of social good that emerges from Big Tech's projects. We then offer data colonialism (Couldry & Mejias, 2019) as a larger theoretical framework in which this emerging form of social good can be understood. When "doing good" is rethought as data colonialism—that is, as a reconfiguration of the social terrain *so that* data can be maximally extracted for economic value—practices that seem benign articulate with more systemic social harms. Data colonialism helps us understand why, to "give," Big Tech must often take away.

Helping the Vulnerable Through Datafication: Critiques and Context

This section reviews critical research on attempts to help socially vulnerable populations through datafied technologies. Such critiques demonstrate clearly the moral failures of these new developments in international philanthropy. Yet, we argue, it can be strengthened through a broader account of the corporatization of social knowledge.

Existing Critiques

A first critical perspective suggests that datafication undermines the basic *rights* of those it would help, continuously tracking and automating life dimensions that would not otherwise be tracked or automated. Well-intentioned technologies and practices of categorization amplify existing vulnerabilities, even when they work well (Jacobsen, 2015). Invasive biometric registration of refugees, for instance, may increase the accountability of humanitarian organizations, but this ignores the potential privacy breaches for unprotected individuals, exposing them to persecution and harassment (Cinnamon, 2020; Madianou, 2019a). A fixation with "innovation" has justified experiments with vulnerable people in emergency contexts (Sandvik, Jacobsen, & McDonald, 2017), which would hardly be tolerated by affluent people in Western countries (Mann, 2018).

These immediate harms cannot be explained by practices and technologies per se: They are rooted in *political economy* dynamics. Discussing humanitarianism, Madianou (2019a) cites the logics of accountability and audit, which control humanitarian organizations and their expenditure; solutionism, which simplifies intricate social contexts; and securitization, which aims to identify and control certain actors deemed "dangerous." Burns (2019) proposes the notion of "philanthro-capitalism" to explain how

[1] The emergence of the welfare state in UK, for example, is connected to the reconfiguration of poverty as a statistical object that could be measured, located, and stratified through surveys (Desrosières, 1998).

humanitarian projects may constitute a "business model and marketing strategy" (p. 15) for private companies. In their study of international development projects in low- and middle-income countries, Taylor and Broeders (2015) note "a combination of datafication and privatization" (p. 230) that weakens local governments, expands markets for corporations (including Big Tech), and deepens inequalities.

Given that these institutional data practices have been commonly developed in richer countries and deployed in poorer nations, many authors see *colonialism* as the best framework to understand them. Some redeploy older definitions of development and humanitarianism as continuing relations of domination between Global South and Global North. On this view, North American and European digital conglomerates resemble modern empires, imposing their culture and values from an unequal flow of capital and data (Anonymous, 2016; Nothias, 2020; Oyedemi, 2020). Madianou (2019b) uses postcolonial concepts to argue that data-powered humanitarian organizations enact a "technocolonialism," "reinvigorate[ing] and rework[ing] colonial relationships of dependency" (p. 2) between South-North. The extraction of data about, for example, refugees is not necessarily aimed at helping refugees, but at justifying "the funding of aid projects" (Madianou, 2019b, p. 8). But this disconnection between humanitarianism and those it allegedly benefits leaves unspecified the precise relation *data* play in the reproduction of colonialism.

The Corporatization of Social Knowledge

The above literature demonstrates how private companies datafy a multitude of social actors and relations in particular fields, notably humanitarianism and international development, often producing harmful consequences. Yet the *corporatization of social knowledge* is by no means confined to these fields. Clarifying this wider process illuminates how social good *everywhere* is being transformed by datafication, opening up new empirical sites to study this transformation.

Data about people today is less a public asset and increasingly privately funded, collected, and analyzed. This process has long historical roots. In parallel to the transformations of social relations required by the advent of industrial capitalism in the 19th century (Polanyi, 1944/2001), there was an important transformation of social knowledge, through state-backed collection of citizens' information and statistical analysis (Desrosières, 1998; Hacking, 1990; Porter, 1996). In the 21st century, a new transformation of social knowledge is underway, driven not by governments but by corporations. The huge increase in commercial knowledge of everyday life since the 1980s now dwarfs what states know about social subjects (Gandy, 1993; Starr & Corson, 1988), a change accelerated by the emergence of commercial platforms (Cohen, 2019). Such transformation empowered new corporate actors to render social life more "trackable and tractable" (Fourcade & Healy, 2017, p. 19). This new model of social governance has fuzzy limits. Once a "social graph" (Farber, 2007, para. 1) is in place, no human interaction seems free from corporate intervention: The very notion of data-driven intervention implies a datafied social good to be actualized. This process, found in both Global North and Global South, is not exclusive to neocolonial conditions.

In this context, the definition of social good *everywhere* becomes anything but innocent, enabling a new domain of privileged social action in which some actors are strong and others are weak (cf. Andrejevic, 2014). "Strong" actors are those (very often corporations rather than governments) that have access to large-scale resources for data collection, storage, analysis, and exchange. "Weak" actors are those over whom strong actors have privileged power to act. Weakness can take many forms: underlying dependence on social services whose provision becomes conditional on those actors entering into relations of data extraction; possession of minimal resources for collecting, storing, and processing data; and greater vulnerability to the outcomes of data-driven judgements.

There are numerous examples of this within the Global North. Supposedly objective data-driven welfare systems are commonly faulty and multiply biased against social minorities, preventing people from accessing benefits and services. As Eubanks (2018) and Gangadharan (2017) demonstrate in the North American context, many individuals are too powerless to stand up to, and perhaps even to become aware of, these injustices. Alston (2019) argues that hybrid "digital welfare" states are being erected "in the name of efficiency . . . individual autonomy, and . . . the imperatives of fiscal consolidation" (p. 4) in which private actors are increasingly important. For governments, datafication is "a Trojan Horse" to further "neoliberal hostility toward welfare and regulation"; for their private partners, discourses about "welfare" help render datafication technologies "benign" (Alston, 2019, pp. 3, 10). These new techniques have thus helped to reduce welfare budgets, narrow the number of benefits and beneficiaries, eliminate services, and impose "stronger sanction regimes" (Alston, 2019, p. 3).

But broader geographical disparities still matter. On a global scale, a further dimension of the dichotomy between strong and weak actors emerges: being a poor and/or vulnerable citizen of *a country* that overall is more vulnerable to net data extraction (Weber, 2017). Particular nation-states are vulnerable to data extraction if they lack large-scale data actors, strong infrastructures for data collection, processing and storage, and have weak infrastructures for connectivity.

The result is what Madden, Gilman, Levy, and Marwick (2017) call a "matrix of vulnerabilities," in which some are disposed to be the targets of data-driven programs for social good, and others the ultimate economic beneficiaries (p. 1). The specific vulnerabilities this matrix generates interconnect with the inequalities inherited from neocolonial geographies, but they are not identical to them. This, we propose, is the process that underlies the datafication of social good: an ever deeper embedding of corporate power in the production of social knowledge, which goes beyond solutionism (Arora, 2019; Morozov, 2013), because it reconfigures the whole domain in which social problems come to be known and need "solutions." A new kind of actor stands out as the privileged beneficiary of such opportunities: global digital technology corporations, which become central empirical sites to understand the relationship between datafication and social good. In the next section, we suggest one entry point to investigate their global role.

Big Tech's "Social Divisions": A Methodology

This section proposes what might be called the "social divisions" of large North American conglomerates of datafication technologies (Big Tech, for short) as a useful—and understudied[2]—empirical entry point. We justify here our methodological decisions and procedures.

[2] But see e.g., Taylor and Broeders (2015).

Owning to their global reach and market dominance, Big Tech's decisions and actions cut across multiple sectors, potentially affecting billions of individuals in radically different contexts and setting standards replicated by countless other organizations. Such large organizations thus provide a transversal perspective for understanding the emerging relationship between social good and datafication. While accepting the difficulty of defining how big a digital tech company must be to be part of the "Big Tech" club, and how exactly size should be measured, our approach is pragmatic. "Big Tech," here, is understood as what public discourse considers "Big Tech" to be. Usually, the term refers to North American global corporations, in particular Google, Apple, Facebook, and Amazon (GAFA), occasionally adding Microsoft (GAFAM); IBM, one of the oldest digital tech conglomerates in the world, is sometimes considered part of the group as well. We *initially* considered all six companies.[3]

Which of these companies' sprawling activities relate to social good? This question poses some problems, since Google and Facebook have long defined themselves as corporations whose goals are fully aligned with those of their users—and, thus presumably, society in general. However, by exploring their corporate websites, we quickly realized that most of these companies appear to have a particular department within their organogram: Big Tech's *social divisions*. By this we mean more or less organized sectors within Google, Facebook, Amazon, Microsoft, and IBM that *define themselves* as geared toward helping (typically, vulnerable) people, not profit. While these divisions do not exhaust those companies' actions linked to social good, their common professed intention renders them comparable. These divisions' names may mimic NGO identities ("Google.org"), be explicit about aims (Microsoft's "Philanthropies"), state who they are directed to (Amazon's "Our Communities"), or simply add the term social good after their brands (Facebook's, IBM's). Google.org seems to be the oldest of its kind, created in 2005; "Facebook Social Good" was launched in 2017; the earliest evidence of its materialization as a website is from 2019 (Facebook, 2019). The others appear to have been created sometime in between. Apple seems unique in that its social good work was not, at the time of the writing, integrated into a single department or website, but fragmented into multiple, smaller projects.

Through systematically reading companies' websites (using Internet Archive's Wayback Machine) and academic and journalistic research on the topic,[4] we reviewed the various projects supported or conducted by Big Tech and their social divisions to identify those related to our focus.[5] We decided to examine 18 specific projects and broader initiatives (see Appendix). Our purposive sampling was guided by our interest in datafication, and its use to supposedly assist vulnerable individuals.[6] While datafication sits at the core of most Big Tech products, their social divisions are not necessarily involved in projects that hinge on datafication. Often, they appear to work as traditional charities, donating resources to other NGOs involved in conventional philanthropic causes (education, climate change, disaster relief). Other

[3] This article does not investigate social good initiatives by Chinese technology conglomerates. Because of their cultural, political, and legal contexts, they are not easily comparable to North American counterparts.
[4] This revealed projects that, while not listed on the "social divisions"' websites, were clearly aligned with these divisions' activities and goals.
[5] Because of our interest in Big Tech companies per se, we did not examine philanthropic organizations founded by firms' owners (e.g., the Chan Zuckerberg Initiative).
[6] Thus, we do not focus on connectivity programs such as Facebook's Free Basics (Nothias, 2020).

projects financed by Big Tech allege to apply datafication for the social good, but their emphasis is not on the *social* world, but exclusively on environmental issues[7]; others do not focus on vulnerable people or only indirectly intend to assist them—so we did not consider them to fit our focus. We did not limit our analysis to initiatives executed only by Big Tech. It sufficed for a project to have a Big Tech company as a partner: Decisions about how to spend resources collaboratively are also important to understand firms' assumptions regarding social good and datafication. Lastly, we were unable to identify a project supported by Apple that was clearly based on datafication. One of its largest programs, "ConnectED," donates devices and assistance to schools and teachers, but does not appear to collect individualized data that can be mined (Apple, 2020; see also Apple, 2019). Thus, this article's conclusions should not be extended to Apple.[8]

Given the notorious opacity in which Big Tech shroud their operations and staff, researching these projects is far from straightforward. Hence, we decided to collect as much public data about our projects sample as possible online. This effort was surprisingly fruitful. The documentation we assembled includes documents made public by companies, files in which we archived the content and screenshots of companies' webpages, legal filings, academic articles, and spreadsheets, where we compiled information—in particular, data about the prizes and winners of Google's Impact Challenge. With the help of Google's patents search engine, we used the names of researchers and companies' employees to look for applications linked with the selected projects, collecting all records we could identify. Then, we conducted thematic analysis of these documents. Initially, we highlighted textual content associated with the ideas about social good on which those projects depended. After this, through an iterative reading of these excerpts, we inductively identified three key assumptions underpinning the analyzed projects, which defined social good as *datafied*, *probabilistic*, and *profitable* (see next section). Treating these three assumptions as our main themes, we reanalyzed some documents to extract further detail.

The data we examined are partial and fragmented, inherently limited by companies' decisions on what to make public, and the extent of our searches. This made it hard for us to be certain about various factual aspects of these projects—for example, all the countries where they have been implemented. Relatedly, we cannot independently confirm that companies' descriptions of what these projects do is accurate: We rely, here, on their self-interested discursive representations. Thus, we do not claim that those documents (or our conclusions) exhaust *all* facets of these projects. Nevertheless, we are convinced that the documentation we collected is substantive enough to ground the unpacked in what follows.

Social Good, According to Big Tech

Our analysis of the selected projects found that, when using datafication techniques aimed at helping vulnerable individuals, Big Tech companies tend to assume social good as *datafied*, *probabilistic,* and *profitable.* Below, we explain these terms through illustrative examples.

---

[7] Social and environmental issues are ultimately inextricable, but our primary focus is the datafication of human beings.
[8] Overall, Apple has been criticized for its scarce philanthropic efforts (Kahney, 2019).

Social Good as Datafied Good

By arguing that social good appears primarily understood as *datafied good* we mean that social good is generally taken as proportional to and made comprehensible by the quantity, type, and granularity of the data that can be gathered. This might seem tautological—projects based on datafication must surely require datafication to be conducted. Indeed, the understanding of social good as datafied good usually remains implicit, a too-obvious-to-mention notion underpinning the very idea of using "data" to "solve the world's toughest problems" (IBM, 2020). But when the imperative to datafy is more explicitly discussed, the tensions buried within that apparent truism emerge.

Take Google.org's report on the hundreds of applicants to its 2018 "AI for Social Good" funding competition. The report combines assessments and summaries of the applications to the "challenge" ("insights") with normative prescriptions ("opportunities"; Google, 2019, pp. 2–3). This provides a rare (if partial and filtered) overview of how multiple actors conceive of the relationship between datafication and social good. While not a definitive institutional account, the report illuminates what social good *ought to be*, per Google. When explaining "data accessibility," the report says that

> access to reliable and meaningful data is a consistent barrier for social sector organizations interested in applying AI methods and capabilities. . . . The data challenges faced by economic empowerment and equality and inclusion proposals illustrate the difficulty in collecting large amounts of data from vulnerable populations that are often more transient, highly sensitive to privacy, and less likely to participate in the formal economy. . . . In sectors where data already exists but is not easily accessible, organizations that own data have an opportunity to invest in data-sharing partnerships and responsible open-sourcing to allow other stakeholders to utilize this data. In these cases, it will be important to consider privacy and security risks as well as potentially harmful use cases before sharing datasets broadly. In more data-sparse sectors, funders can help finance data collection. Funders and policymakers could leverage their resources and influence to support the collection and sharing of data, where appropriate. (Google, 2019, pp. 16–17)

Here, dearth of data is framed as a "barrier" that endangers the very feasibility of social good. The imperative of getting hold of data involves different strategies ("collection," "sharing," "open-sourcing"), and even justifies the call for "funders and policymakers" to use their "resources and influence" to construct data sets about "vulnerable" populations (Google, 2019, p. 17).

Obliquely recognizing the controversial nature of such aggressive recommendation, the report reminds us that it is important to consider "privacy and security risks," presumably for those "vulnerable" populations who are "often more transient, highly sensitive to privacy, and less likely to participate in the formal economy" (Google, 2019, p. 16). Yet such care seems only perfunctory, given the absence of any hint on what sorts of concrete limits should be imposed on the monitoring of vulnerable people and through what gradations of privacy. If some people are "highly sensitive to privacy," (Google, 2019, p. 16) it follows that others are less sensitive or even not sensitive at all. In suggesting there are two categories of rights holders, the report contradicts a foundational idea of modern privacy—universality. There is no acknowledgment of individuals' ability to discuss and understand those possible breaches, nor of the unequal dynamics that shape such ability. It is up to technologists and social entrepreneurs to decide what counts as harmful and how to avoid it. What those who they allegedly seek to help and protect think about their own well-being is not discussed.

Often, in the projects reviewed, it is not just vulnerable people who need to be datafied for their good to be realized. Consider Facebook's "Social Good" portal, much of which is dedicated to explain how nonprofits can create pages on the social media platform to grow "its community of supporters and create more connections and interactions with people" (Facebook, 2020b). The portal provides a "best practices" primer on how to do this. Its guiding principle is, the more data in relation to a nonprofit is created by the organization *within Facebook*, the likelier it is for the nonprofit to be successful and social good to be done. More data means that Facebook can better profile the organization and its potential donors, and is more likely to connect them in an efficient manner: "When people share interests and ideas on Facebook, it helps you find and connect with those who care most about your work" (Facebook, 2020b). But the *social* reasoning behind this self-serving logic is not made clear. At best, it is an assumption that personal life is naturally there to be marketed. A page called "Marketing 101 for Non-Profits" suggests to "share relevant personal stories from members of your organization that showcase their experiences . . . personal stories from your staff, supporters and beneficiaries that may inspire people to share their own" (Facebook, 2020a). The possibility that individuals might not want to publicize their lives to their employer is ignored.

Nonprofits are also repeatedly reminded of the value entailed in using Facebook's data tools. Organizations are encouraged to "know" and "target" the audience of their posts by "demographics" and "interests." We can see this as part of everyday data practices in a datafied society (Kennedy, 2016), but it is the underlying assumption in which we are interested here: that doing social good is reducible to counting and parsing interactions already datafied by the platform. The novelty is not that NGOs need to relentlessly market themselves, but that the successful marketing of their (presumably diverse) goals is assumed to be reliant on datafication processes controlled by one company—Facebook.

Social Good as Probabilistic Good

The idea of *probabilistic good* also flows from the conflation of datafication and any attempt to help vulnerable people. Datafication hinges on the need to transform all dimensions of life into data, but also on making sense of this data through predictive computational systems whose language is often probability (Domingos, 2015). Our point is not simply to say that Big Tech employs or supports social good projects based on probabilistic systems (what is far from novel), but to argue that once social good is datafied, its realization begins to be understood as *necessarily* probabilistic. As with datafied good, the assumption of probabilistic good is usually taken for granted, suggested in the vague but recurrent references to "AI" and "machine learning"; when, however, probability is explicitly debated, its complex association with this conception of social good becomes easier to gauge.

A useful example is Facebook's "Suicidal Prevention Tool," through which the platform identifies posts about self-harm, have them reviewed by human moderators, and show "support options, such as prompts to reach out to a friend and help-line phone numbers" (Card, 2018, para. 9) to original posters.[9] An ampler view is offered in a patent application that Facebook filed to claim intellectual rights over the system. When describing how the tool works, the application says how "background" and interactional data can "form a pattern or fingerprint which may be used to *infer* whether a set of behaviours associated with the user(s) is indicative of a likelihood of suicide and self-injury" (Muriello, Ben-David, Guadagno, Callison-Burch, & Tauro, 2019, p. 4). At another moment, the application explains that the "probability" of a user engaging in self-harm might be an "output" of the tool, representable by a "set of values" indicating "that a user may require self-risk injury mitigation, a confidence level that a value of [a] classification [of the user as at risk or not] is correct, a severity level corresponding to [this] classification" (Muriello et al., 2019, p. 6). That is, an analysis of different forms of data (including "background" data; i.e., data that are not related to the post itself) will never say whether a post *is indeed* a signal that someone is at risk of suicide and should be helped. It only displays the probability that this *might* be the case.

The ethics of designing a probabilistic tool about a social act as serious as suicide are pondered in an academic paper penned by Facebook employees, in which they say,

> If we wanted to ensure we caught every single post expressing suicidal intent then we would want to review every post put on Facebook, but of course that is impossible. ML [machine learning] is probabilistic in nature so it will never be possible to ensure 100% of accuracy in its use. . . . How can we target the relevant posts and allocate the strictly necessary resources for that, while being as thorough as we can? (de Andrade, Pawson, Muriello, Donahue, & Guadagno, 2018, p. 681)

In machine learning terms, they argue, "This is a question of how to set the threshold" (de Andrade et al., 2018, p. 681):

> If we lower the threshold, the more posts that will less likely be actionable will need to be reviewed by more people; this poses the risk of having a disproportionate number of human reviewers looking at non-concerning posts. If we raise the threshold, the more accurate will these posts be and the fewer people we will need to do the human review of the content; but this runs the risk of missing content that should have been flagged and reviewed. In response to this challenge, our philosophy has been to maximize the use of human review available to us without falling beneath a certain threshold of accuracy. We have thus substantially increased our staffing in this area. (de Andrade et al., 2018, p. 682)

The authors are candid: A machine-learning-powered tool will never be fully accurate, and decisions on how to make the tool more efficient will take into consideration elements that go beyond social good, such as "resources," defined here as how many people should be hired to make decisions on posts. Their decision is ethically defensible: instead of missing more posts that could be potentially about self-harm, they hired more people, even if this made the tool more expensive. Many users likely benefited from the

[9] See also Ananny (2019) for an analysis of the probabilistic nature of this project.

tool. But how many suicidal posts were missed by this system? It is hard not to understand such a "trade-off" as, ultimately, a probabilistic damage control.

Facebook's decision to do social good probabilistically entailed necessarily its opposite—*some* likelihood of *harm*. For sure, no attempt to help vulnerable others can be expected to succeed perfectly. What is new here is the a priori *presumption* that this social good project *will* fail in a percentage of cases, not due to a lack of resources but because, with machine learning, accuracy (as the basis for doing good) is always, and "naturally," probabilistic. Yes, it is always better to save some lives instead of not saving any life at all. However, by contrast with other large-scale phenomena, whose complexity is so great that probabilistic reasoning comprises the only possible option (e.g., pandemics), suicide is a fairly well understood problem from the beginnings of sociology as a discipline (Hacking, 1990). There is a different trade-off operating here, hardly explored in the Facebook employees' paper, between using datafication and using other, more traditional and efficient forms of mental health support. Facebook could have arguably saved more lives if, instead of employing machine learning, it had focused on promoting and funding local suicide prevention hotlines, for instance (see Miller, 2019).

A further example of probabilistic social good comes from Project Horus, a "collaboration" between Microsoft and the government of the Argentinean province of Salta "to apply artificial intelligence in the prevention of teenage pregnancy and school dropout [rates]" (Microsoft, 2018, para. 10).[10] The project relied on the "permanent" monitoring of the habits and the bodies of poor women and children with the goal of constructing "complete knowledge" about them (Abeleira, 2018, pp. 46, 71). This data was then analyzed by "smart algorithms," which could "allow [the project] to identify characteristics that could lead to one of these problems [teenage pregnancy and school dropout rates] and warn the government so that they could work on their prevention" (Microsoft, 2018, para. 10).

The project has been criticized for its glorification of total surveillance of vulnerable individuals, its association with antiabortion movements, and its technical errors (Peña & Varon, 2019). Even without these issues, the ambiguities of how it understands social good probabilistically would remain. As one of the creators of the project said, "The model we developed has an accuracy level of almost 90% from a pilot test" (Microsoft, 2018, para. 11). What about the other 10%—on what basis can one argue that they do not deserve help from government? Again, it is hard to see why such new approach should replace traditional policies (e.g., proper schools, well-trained and well-paid teachers, universal access to contraceptive methods).

Probabilistic good defies intuitive perceptions of what "good" means and how it can be instantiated. Not because it involves some attention to probability—long-standing utilitarian approaches to ethics do also, and most individuals calculate probabilities of certain outcomes when making everyday decisions, but always on the basis of assuming in advance *which* would be better of various possible outcomes. The problem rather, in Big Tech's approach, is that the association between (probably) doing good and (necessarily) allowing some harm to happen flows automatically from the probabilistic notion of algorithmic knowledge on which proponents of datafied social good *choose to* rely, as their model for producing social knowledge.

[10] Similar "collaborations" were initiated in Brazil (Ministério do Desenvolvimento Social, 2017) and India (Rao, 2018).

In opting to deploy machine learning systems, they make inevitable and, thus, acceptable that not all people who deserve to be helped will indeed be helped. Social good is stripped of its universality as a goal, becoming at best the orientation of a probabilistic calculation process.

Social Good as Profitable Good

As if acknowledging the contradictory relationship between capitalism and ethics, Big Tech's social divisions are keen to define themselves as disconnected from companies' business models. Yet on closer inspection, the projects they conduct and support seem inseparable from the main goal of these firms—to generate profit. The social good that datafication can do is thus expected by Big Tech also to be a form of *profitable good*. This assumption echoes one of the oldest and most common criticisms of charitable organizations—namely, that they represent capitalism in disguise (McGoey, 2015). What sets Big Tech's practices apart is *how* they create economic value directly out of the act of doing social good.

Let us begin with what seems universal to all projects analyzed: their role in Big Tech's marketing strategies. This is evidenced by the decision to both give visibility to these initiatives and to describe them in a way that constantly defines companies' identity as driven by moral goals. More systemically, this marketing strategy might be understood as not only trying to elevate companies' moral status but also sanitizing the methods on which their business model depends. If datafication can save the world, why worry about its hazards? When Big Tech's social good projects are used as PR instruments, we remember that, ultimately, they serve economic (not "social") goals. But as such, marketing does not directly generate income. The same cannot be said of a different set of practices, whereby social good projects are *materially* associated with each corporation's business model.

One such practice is the entanglement of social good projects with commercial products. This seems particularly true for Facebook, since most of its social good projects seem to be built on top of its platforms and are often subject to the similar technological and legal arrangements applied to other actions that happen in these platforms. An example is the "Donate Button," a donation system that collects "card numbers and other payment method information, and information such as your transaction history or a copy of your ID," which might be used for "our (or others') legitimate interests, including our interests in providing an innovative, personalised, safe and profitable service to our users and partners" (Facebook, 2020c, para. 6, 14,). As this excerpt of the tool's privacy policy makes clear, the platform might use the data associated with a donation for commercial profiling and targeting purposes.

Something similar seems to have been done by Google through its G Suite for Education, often supplied *at no cost* to schools in various countries (Google, 2020). The Suite combines many of the company's leading software services (Gmail, Calendar, Drive, Docs, Sheets) and, eventually, its low-cost laptop (Chromebook). However, according to the Attorney General for the State of New Mexico (2020), in the U.S., G Suite was used "to collect large quantities of valuable personal information, without their parents' consent, from children under 13 who are often required by their schools to use these services" (p. 1). According to a complaint filed by the attorney, G Suite collected students "physical locations; websites they visit; every search term they use in Google's search engine (and the results they click on); the videos they watch on YouTube; personal contact lists; voice recordings; saved passwords; and other behavioral information" (Attorney General for the State of New Mexico, 2020, p. 5). More than that, it "mined students' email accounts" and "used that data . . . for advertising purposes" (Attorney General for the State of New Mexico, 2020, p. 5). To the press, the company called the claims "factually wrong," but did not explicitly deny that it had collected data (Statt, 2020, para. 9).

If in both Facebook's and Google's cases what's at stake is collecting as much data as possible, other projects tie doing social good to initiatives that enhance the buying of their products. Consider Amazon's "Alexa Skills Challenge: Tech for Good," which in 2018 "invited developers" to build apps ("skills") for the company's virtual assistant technology "that would have a positive impact on the environment, local communities, and the world" (Vacherot, 2018, para. 1). That is, to get access to any of the apps, people would have to first acquire or at very least use Alexa.

Sometimes, the process of developing social good tools and practices might engender potentially profitable nondata assets—more specifically, patents. IBM has been for decades the U.S. leader in granted patents, and considers intellectual property a key part of its business model (IBM, 2019). The company boasts that "while working to solve some of the toughest challenges facing our world, novel solutions resulted in 9 pending patents" (IBM, 2020). One such application regards a method to collect digital data about a humanitarian crisis (Soares et al., 2017). Facebook is another Big Tech company that has used social good projects to invent patentable technologies. Above, we cited the patent of its "suicide prevention tool," but the company has applied for (and sometimes been granted) the rights over systems concerning, for instance, donations (Subbarayan, Agarwalla, Triolo, & Quirino, 2015). As a 2019 corporate blog post on Facebook's "approach to patents" makes clear, developing patents is always part of Facebook's attempt to gain "market advantage" (Chan, 2019, para. 3, 6). More important than the question of *how* profitable such patents will be is the underpinning assumption that social good technologies—often developed thanks to the data of unaware individuals—are part of a wider enterprise, one of whose core goals is to generate enforceable and sellable (thus profitable) rights to private property.

Social Good and the Project of Data Colonialism

Our argument so far has been that not only are Big Tech companies actively involved in using datafication for social good, but that this involvement achieves another hidden and more consequential goal: the progressive reconfiguration of the social domain *itself*, or at least ever larger parts of it, in ways that position those Big Tech companies as privileged providers of *social solutions* and privileged purveyors of *social knowledge.* The social solutions and knowledge that Big Tech companies provide have three features: They are, first, datafied; second, by being datafied, they are often probabilistic; and third, since they are the output of large commercial corporations, they aim to be profitable. This new commercially driven production of the social good is at an early stage of its unfolding, but already it signals a profound rebalancing of power and governance in the domain of social life, privileging corporations with large-scale data power and making states (and other commercial and civil society actors) dependent on those corporations. The result is more than digital solutionism: It is a refashioning of the tools of social intervention so that a particular kind of digital solutionism *necessarily* seems the only toolkit available.

The social good is not a neutral fact, but a set of socially constructed parameters by reference to which good and consequential actions in the territories we share are evaluated. A corporatization of the social good has consequences for how social life is known, understood, and governed. We now contextualize this process within data colonialism (Couldry & Mejias, 2019) and specifically data colonialism's deepest continuity with historic colonialism: a shared conception of *rationality* that reduces the human lifeworld via a single way of reading the actual "heterogeneity of all reality" (Quijano, 2007, p. 177). The proposition that Big Tech, based in one part of the world and benefiting from a very particular concentration of resources, can judge how social problems should be interpreted and resolved across *all* the world's societies is, in the light of colonial history, an astonishing usurpation of power that claims the capacity to see all the world's social differences and similarities in terms of one single data-driven logic that justifies corporate intervention anywhere. That is why data-driven solutions are rarely offered as part of a *range* of solutions to social problems, but as *the* solution, an expression of a new language for defining and solving social problems that replaces all others. No one is asked if they agree with this act of substitution of social knowledge, or its consequences in terms of data collection and processing. In the process, populations' freedom to define *their* social good, *their* version of social knowledge, is overridden.

The data colonialism thesis provides a larger framework in which to grasp the datafication of social good. To recap, the data colonialism thesis is the proposal that what is happening with data today across the world constitutes a genuinely new stage of colonialism: an epochal act of resource extraction that bears comparison with the original territorial "landgrab" (Dörre, Lessenich, & Rosa, 2015, pp. vii–viii) by European powers under historic colonialism. This new landgrab targets not physical land and the resources that flow from it, but human life itself, annexing it for capital through technologies of data extraction (Couldry & Mejias, 2019).

One specific aspect of the data colonialism thesis is particularly important for our argument here. This is the idea that, as it develops, data colonialism transforms not just our relations with digital interfaces but also the very ground on and from which social knowledge is produced:

> Under data colonialism . . . capitalism begins to imagine away any outside to the economy. Its distinctive forms of social knowledge describe a social world that is literally coextensive with economic life. . . . Instead of "social relations [being] embedded in the economic system" [Polanyi, 1944/2001, p. 60] . . . , social relations become the economic system, or last a crucial part of it, as human life is converted into raw material for capital via data. (Couldry & Mejias, 2019, p. 117)

This flows from the double nature of data as both economic value *and* a source of potential knowledge. As a result, claims about what should *be done by humans* in the social world automatically become claims about *what data* should do in that world, generating *future value* for the very entities that will lead or aid the production of such data. In this way, social good projects extend further the colonizing impulse of digital platforms to "produce 'the social' for capital" (Couldry & Mejias, 2019, p. 26), not just (as do digital platforms) inciting social activity from which the production of value through data can be optimized, but reshaping social life as a whole—and the tools of governance that seek to manage it as a whole—around the ever-increasing production of data. Levels of data production become *themselves* an index of social good. In this way, data-rich corporations are installed as privileged reference points not only for convening the social present (Facebook's vision of its global "community") but also for shaping and governing the social future (the idea of "Tech for Good"). For that reason, data-driven projects oriented to the social good are consequential whether they do direct social harm or not, and whether they intervene directly in the social terrain or, as with Google's Impact Challenge, evaluate others' attempts to "do good in the world."

The rationality of data colonialism does not recognize a social world where it makes sense to consult vulnerable people about how new social knowledges are generated; rather Big Tech seeks to act directly *on* the datafied world that it sees and measures. What matters therefore is not whether such projects are done well or badly (Floridi, Cowls, King, & Taddeo, 2020), but that they are done *at all*. Simply by being implemented such projects create new social domains into which data colonialism can further expand, potentially transforming the relations of states to the territories they govern (Magalhães & Couldry, 2020) and "subordinat[ing] considerations of human well-being and human self-determination to the priorities and values of powerful economic actors" (Cohen, 2019, p. 73).

Conclusion

Colonialism has never been *only* about the relentless exploitation of resources for economic gain. Enmeshed in the project to rob and dominate through violence and force was the project to transform the knowledge and governability of the territories called "colonies" and allegedly "civilize" their populations.

When Big Tech and their social divisions deploy datafication to do social good, here, too, economic and moral ideals feed each other—but differently. For, as we have demonstrated, helping vulnerable people (something unobjectionable) becomes *itself* a site of exploitation. This contradiction permeates all the projects that we analyzed. The realization of a *datafied*, *probabilistic*, and *profitable* social good depends on the imposition of certain unfreedoms (in the realm of social knowledge and individual subjectivity) as the cost of apparently protecting certain freedoms (from vulnerability to particular social harms). Such projects rely on an unsaid denial of individuals' capacity to define what "good" ought to be, and so extend the project of data colonialism.

More than arguing that datafication can sometimes produce the opposite of good, a point already clear in the literature reviewed in the beginning of the article, we have discussed social harms that are more than an accident caused by Big Tech's malfunction or inattention. Such harms are an intrinsic part of how these companies operate, a natural consequence of the rationality—at heart a colonial rationality—now applied within both Global North *and* Global South, that underlies their business model. As such, no ethical guidelines can ensure that social good will truly be realized through such projects of datafication. This will happen only if the *whole project* of solving social problems through large-scale data processing concentrated in, or enabled by, large technological corporations is reviewed in the light of colonialism's deep and long-standing entanglements of power and knowledge. Our article has, we hope, contributed to that reassessment.

References

Abeleira, C. (2018). *Project Horus*. Salta, Argentina: Consejo Federal de Inversiones.

Alston, P. (2019, October 11). *Report of the Special Rapporteur on extreme poverty and human rights*. Retrieved from https://undocs.org/A/74/493

Ananny, M. (2019, August 21). *Probably speech, maybe free: Toward a probabilistic understanding of online expression and platform governance*. Retrieved from https://knightcolumbia.org/content/probably-speech-maybe-free-toward-a-probabilistic-understanding-of-online-expression-and-platform-governance

Andrejevic, M. (2014). Big Data, big questions: The Big Data divide. *International Journal of Communication, 8*(17), 1674–1689.

Anonymous. (2016, November 15). Data colonialism: Critiquing consent and control in "tech for social change." *Model View Culture*, *43*. Retrieved from https://modelviewculture.com/pieces/data-colonialism-critiquing-consent-and-control-in-tech-for-social-change

Apple. (2019, January 21). *Out of a culture of giving, a world of difference*. Retrieved from https://www.apple.com/newsroom/2019/01/out-of-a-culture-of-giving-a-world-of-difference/

Apple. (2020). *Apple and ConnectED*. Retrieved from https://www.apple.com/connectED/

Arora, P. (2019). *The next billion users*. Cambridge, MA: Harvard University Press.

Attorney General for the State of New Mexico. (2020, February 20). *Attorney General Balderas Sues Google for Illegally Collecting Personal Data of New Mexican School Children*. Retrieved from https://www.nmag.gov/uploads/PressRelease/48737699ae174b30ac51a7eb286e661f/AG_Balderas_Sues_Google_for_Illegally_Collecting_Personal_Data_of_New_Mexican_School_Children.pdf

Burns, R. (2019). New frontiers of philanthro-capitalism: Digital technologies and humanitarianism. *Antipode*, *51*(4), 1101–1122. doi:10.1111/anti.12534

Card, C. (2018, September 10). *How Facebook AI helps suicide prevention*. Retrieved from https://about.fb.com/news/2018/09/inside-feed-suicide-prevention-and-ai/

Chan, J. (2019, August 29). *How patents drive innovation at Facebook*. Retrieved from https://about.fb.com/news/2019/08/how-patents-drive-innovation/

Chui, M., Harryson, M., Manyika, J., Roberts, R., Chung, R., van Heteren, A., & Nel, P. (2018). *Notes from the AI frontier*. New York, NY: McKinsey & Company.

Cinnamon, J. (2020). Data inequalities and why they matter for development. *Information Technology for Development*, *26*(2), 214–233. doi:10.1080/02681102.2019.1650244

Cohen, J. E. (2019). *Between truth and power.* Oxford, UK: Oxford University Press.

Couldry, N., & Mejias, U. A. (2019). *The costs of connection*. Stanford, CA: Stanford University Press.

de Andrade, N. N. G., Pawson, D., Muriello, D., Donahue, L., & Guadagno, J. (2018). Ethics and artificial intelligence: Suicide prevention on Facebook. *Philosophy & Technology*, *31*(4), 669–684. doi:10.1007/s13347-018-0336-0

Desrosières, A. (1998). *The politics of large numbers*. Cambridge, MA: Harvard University Press.

Dijck, J. van. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208. doi:10.24908/ss.v12i2.4776

Domingos, P. (2015). *The master algorithm*. New York, NY: Basic Books.

Dörre, K., Lessenich, S., & Rosa, H. (2015). *Sociology, capitalism, critique*. London, UK: Verso.

Eubanks, V. (2018). *Automating inequality*. New York, NY: St. Martin's.

Facebook. (2019, June 7). *Social good*. Retrieved from https://web.archive.org/web/20190607220253/https://socialgood.fb.com/

Facebook. (2020a). *Connect with supporters on your Facebook page to strengthen your relationships*. Retrieved from https://socialgood.fb.com/learning-support/connect-with-your-audience/en-GB/

Facebook. (2020b). *Create a page for your charity.* Retrieved from https://socialgood.fb.com/learning-support/getting-started/create-a-page-for-your-nonprofit/en-GB

Facebook. (2020c). *Facebook international limited privacy policy*. Retrieved from https://www.facebook.com/payments_terms/EU_privacy

Farber, D. (2007, May 24). Facebook's Zuckerberg uncorks the social graph. *ZDnet*. Retrieved from https://www.zdnet.com/article/facebooks-zuckerberg-uncorks-the-social-graph/

Floridi, F., Cowls, 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. doi:10.1007/s11948-020-00213-5

Fourcade, M., & Healey, K. (2017). Seeing like a market. *Socio-Economic Review, 15*(1), 9–29.

Gandy, O. (1993). *The panoptic sort*. Boulder, CO: Westview.

Gangadharan, S. P. (2017). The downside of digital inclusion: Expectations and experiences of privacy and surveillance among marginal Internet users. *New Media & Society, 19*(4), 597–615. doi:10.1177/1461444815614053

Google. (2019, September 10). *Accelerating social good with artificial intelligence*. Retrieved from https://services.google.com/fh/files/misc/accelerating_social_good_with_artificial_intelligence_google_ai_impact_challenge.pdf

Google. (2020). *Google for education*. Retrieved from https://edu.google.com/?modal_active=none

Hacking, I. (1990). *The taming of chance*. Cambridge, UK: Cambridge University Press.

Hao, K. (2020, April 23). *Doctors are using AI to triage COVID-19 patients.* Retrieved from https://www.technologyreview.com/2020/04/23/1000410/ai-triage-covid-19-patients-health-care/

Hilbert, M. (2016). Big data for development: A review of promises and challenges. *Development Policy Review*, *34*(1), 135–174. doi:10.1111/dpr

IBM. (2019). *Annual report 2019.* Retrieved from https://www.ibm.com/annualreport/assets/downloads/IBM_Annual_Report_2019.pdf

IBM. (2020). *Science for social good*. Retrieved from https://www.research.ibm.com/science-for-social-good/

International Telecommunication Union. (2020). *AI for good: About us*. Retrieved from https://aiforgood.itu.int/about-us/

Jacobsen, K. L. (2015). *The politics of humanitarian technology.* London, UK: Routledge.

Kahney, L. (2019, April 17). *Is Apple's newfound sense of charity anything other than a marketing ploy?* Retrieved from https://tech.newstatesman.com/business/apple-newfound-charity-marketing-ploy

Kennedy, H. (2016). *Post, mine, repeat*. London, UK: Palgrave Macmillan.

Lewis, P., Conn, D., & Pegg, D. (2020, April 12). UK government using confidential patient data in coronavirus response. *The Guardian*. Retrieved from https://www.theguardian.com/world/2020/apr/12/uk-government-using-confidential-patient-data-in-coronavirus-response

Madden, M., Gilman, M., Levy, K., & Marwick, A. (2017). Privacy, poverty, and big data: A matrix of vulnerabilities for poor Americans. *Washington University Law Review, 95*(1), 53–125.

Madianou, M. (2019a). The biometric assemblage: Surveillance, experimentation, profit and the measuring of refugee bodies. *Television and New Media*, *20*(6), 581–599. doi:10.1177/1527476419857682

Madianou, M. (2019b). Technocolonialism: Digital innovation and data practices in the humanitarian response to refugee crises. *Social Media + Society*, *5*(3), 1–13. doi:10.1177/2056305119863146

Magalhães, J. C., & Couldry, N. (2020, April 24). *Tech giants are using this crisis to colonize the welfare system*. Retrieved from https://www.jacobinmag.com/2020/04/tech-giants-coronavirus-pandemic-welfare-surveillance

Mann, L. (2018). Left to other peoples' devices? A political economy perspective on the big data revolution in development. *Development and Change*, *49*(1), 3–36. doi:10.1111/dech.12347

McGoey, L. (2015). *No such thing as a free gift*. London, UK: Verso.

Meier, P. (2015). *Digital humanitarians*. Boca Raton, FL: CRC.

Microsoft. (2018, April 2). *Avanza el uso de la inteligencia artificial en la Argentina con experiencias en el sector público, privado y ONGs* [The use of artificial intelligence advances in Argentina with experiences in the public, private and NGO sectors]. Retrieved from https://news.microsoft.com/es-xl/avanza-el-uso-de-la-inteligencia-artificial-en-la-argentina-con-experiencias-en-el-sector-publico-privado-y-ongs/

Miller, G. (2019, August 29). *Three suicide prevention strategies show real promise. How can they reach more people?* Retrieved from https://www.sciencemag.org/news/2019/08/three-suicide-prevention-strategies-show-real-promise-how-can-they-reach-more-people

Ministério do Desenvolvimento Social. (2017, September 9). *Parceria entre governo brasileiro, província argentina e Microsoft irá ajudar no monitoramento do Criança Feliz* [Partnership between Brazilian government, Argentinian province and Microsoft will help with the monitoring of Happy Child program]. Retrieved from http://mds.gov.br/area-de-imprensa/noticias/2019/setembro/parceria-entre-governo-brasileiro-provincia-argentina-e-microsoft-ira-ajudar-no-monitoramento-do-crianca-feliz

Morozov, E. (2013). *To save everything, click here*. New York, NY: PublicAffairs.

Muriello, D., Ben-David, D. M., Guadagno, J. L., Callison-Burch, V., & Tauro, S. R. (2019). *U.S. Patent Application No. 15/667,543*. Washington, DC: U.S. Patent and Trademark Office.

Newton, C. (2020, May 2). *Why countries keep bowing to Apple and Google's contact tracing app requirements*. Retrieved from https://www.theverge.com/interface/2020/5/8/21250744/apple-google-contact-tracing-england-germany-exposure-notification-india-privacy

Nothias, T. (2020). Access granted: Facebook's free basics in Africa. *Media, Culture & Society*, *42*(3), 329–348. doi:10.1177/0163443719890530

Oyedemi, T. D. (2020). Digital coloniality and "Next Billion Users": The political economy of Google Station in Nigeria. *Information, Communication & Society*. Advance online publication. doi:10.1080/1369118X.2020.1804982

Peña, P., & Varon, J. (2019). Decolonising AI: A transfeminist approach to data and social justice. *Global Information Society Watch 2019*. Johannesburg, South Africa: Association for Progressive Communications.

Polanyi, K. (2001). *The great transformation*. Boston, MA: Beacon. (Original work published 1944)

Porter, T. (1996). *Trust in numbers.* Princeton, NJ: Princeton University Press.

Quijano, A. (2007). Coloniality and modernity/rationality. *Cultural Studies 21*(2/3), 168–178.

Rao, U. (2018). Govt ties up with Microsoft to check dropouts. *Times of India*. Retrieved from https://timesofindia.indiatimes.com/city/visakhapatnam/govt-ties-up-with-microsoft-to-check-dropouts/articleshow/63863010.cms

Sandvik, K. B., Jacobsen, K. L., & McDonald, S. M. (2017). Do no harm: A taxonomy of the challenges of humanitarian experimentation. *International Review of the Red Cross*, *99*(904), 319–344.

Shi, Z. R., Wang, C., & Fang, F. (2020). *Artificial intelligence for social good: A survey*. Retrieved from https://arxiv.org/pdf/2001.01818.pdf

Soares, I. M. B., Dhurandhar, A., Kumar, A., Mojsilovic, A., Pham, K. T., Varshney, K. R., & Vukovic, M. (2017). *U.S. Patent Application No. 15/484,325*. Washington, DC: U.S. Patent and Trademark Office.

Starr, P., & Corson, R. (1988). Who will have the numbers? The rise of the statistical services industry and the politics of public data. In W. Alonso & P. Starr (Eds.), *The politics of numbers* (pp. 415–447). New York, NY: Russell Sage Foundation.

Statt, N. (2020). *Google sued by New Mexico attorney general for collecting student data through Chromebooks*. Retrieved from https://www.theverge.com/2020/2/20/21145698/google-student-privacy-lawsuit-education-schools-chromebooks-new-mexico-balderas

Subbarayan, A., Agarwalla, B. K., Triolo, C. J., & Quirino, T. M. (2015). *U.S. Patent Application No. 14/589,956*. Washington, DC: U.S. Patent and Trademark Office.

Taylor, L., & Broeders, D. (2015). In the name of development: Power, profit and the datafication of the Global South. *Geoforum*, *64*, 229–237. doi:10.1016/j.geoforum.2015.07.002

Vacherot, B. (2018). *Announcing the winners of the Alexa Skills Challenge: Tech for good*. Retrieved from https://developer.amazon.com/blogs/alexa/post/ca34b954-1c5d-4a59-b326-f45c8df7c89c/alexa-skill-tech-for-good-challenge-winners

Weber, S. (2017). Data, development, and growth. *Business and Politics*, *19*(3), 397–423. doi:10.1017/bap.2017.3

Appendix: Initiatives and Projects Analyzed

Google
- Google AI Impact Challenge 2018
- Google for Education
- Google for Nonprofits
- Chance
- Bayes Impact

Facebook
- Charitable Giving
- Crisis Response
- Health
- Mentorship

IBM
- Prescription Guidelines for Opioid Epidemic
- Neurology-as-a-Service
- Cognitive Financial Advisor for Low-Wage Workers
- Causal Pathways Out of Poverty

Microsoft
- Project Horus
- AI for Health
- AI for Accessibility
- AI for Humanitarian Action

Amazon
- Alexa Skills Challenge: Tech for Good