Original Research Article

BIG DATA & SOCIETY

Big Data & Society
January–June: 1–12

DOI: 10.1177/20539517231173901
journals.sagepub.com/home/bds
Sage

Stepping back from Data and AI for Good – current trends and ways forward

Ville Aula[1] and James Bowles[2]

Abstract

Various 'Data for Good' and 'AI for Good' initiatives have emerged in recent years to promote and organise efforts to use new computational techniques to solve societal problems. The initiatives exercise ongoing influence on how the capabilities of computational techniques are understood as vehicles of social and political change. This paper analyses the development of the initiatives from a rhetorical slogan into a research program that understands itself as a 'field' of applications. It discusses recent academic literature on the topic to show a problematic entanglement between the promotion of initiatives and prescriptions of what 'good' ought to be. In contrast, we call researchers to take a practical and analytical step back. The paper provides a framework for future research by calling for descriptive research on the composition of the initiatives and critical research that draws from broader social science debates on computational techniques. The empirical part of the paper provides first steps towards this direction by positioning Data and AI for Good initiatives as part of a single continuum and situating it within a historical trajectory that has its immediate precursor in ICT for Development initiatives.

Keywords

Data for good, AI for social good, data science for social good, ICT for development, critical data studies, technological initiatives

Introduction

Past years have seen a significant increase in attempts to tackle grand challenges with new computational tools. In this article, we examine the consolidation of 'Data for Good' and 'AI for Good' as initiatives for the promotion of computational technology as a solution to societal problems.

We suggest that Data and AI for Good[1] are more than just pithy slogans or coincidental connections. Instead, the aspirations, publicity, networks and resources associated with Data and AI for Good shape how researchers and practitioners understand the potential computational technologies. Between 2014 and 2020, Data and AI for Good were invented, popularised and established as a loose yet coordinated scientific, corporate and practical program.

The paper draws together the history of, and emerging literature on, Data and AI for Good initiatives to offer a critical analysis of the phenomenon. The paper provides an overview of key initiatives, highlighting their continuity and conceptual links with the field of ICT for Development. We present a framework for different analytical orientations towards the initiatives, contrasting between prescriptive and descriptive studies and between a focus on initiatives themselves and underlying theoretical themes. The paper, therefore, does not aim to be empirically exhaustive but use original empirical analysis to consolidate earlier findings and establish empirical and theoretical vistas that are generative of further critical research.

Specifically, the paper expands on initial critical analyses by Madianou (2021), Magalhães and Couldry (2021) and Holzmeyer (2021). It also connects to the critical appraisal of ethical thinking in technology start-ups that follow the slogan 'Tech for Good' (Powell et al., 2022). These works offered critical readings of individual projects or programs but did not attempt to bring together a broader picture of Data and AI for Good as a family of initiatives, which is the goal of this paper. The timing of our intervention is important. Although first instances of promoting computational technologies 'for Good' were arguably a rhetorical strategy, the initiatives have later successfully established themselves as a specific 'field' whose

[1]Department of Media and Communications, London School of Economics and Political Science, London, UK
[2]Third Sector Research Centre, School of Social Policy, University of Birmingham, Birmingham, UK

Corresponding author:
Ville Aula, Department of Media and Communications, London School of Economics and Political Science, London, UK.
Email: v.v.aula@lse.ac.uk

importance is promoted in programmatic journal articles (e.g. Tomašev et al., 2020) and whose technical achievements are surveyed in systematic reviews of the emerging 'field' (e.g. Shi et al., 2020). The analysis is therefore timely both because of the newly developed maturity of Data and AI for Good initiatives, and because enough literature has emerged to make sense of them.

We argue that researchers should take both a practical and an analytical step back from the initiatives. A practical step back means that computational researchers should be cautious of uncritically promoting the initiatives. The analytical step back means that researchers inclined towards social sciences and critical analysis should concentrate on the bigger picture of related initiatives, examining the similarities and differences of past and present movements and broader theoretical themes that underlie them. Overall, we argue *against* Data and AI for Good initiatives requiring their own specialist literature, as we have seen with ICT for Development, because this would only exacerbate discontinuity between related academic fields when the need is to build a unified body of critical literature on computational technologies.

The paper is organised as follows. First, we offer an outline of how different Data and AI for Good initiatives have emerged and how they link together. In the second and third sections, we identify links between the current initiatives and ICT for Development. After setting these empirical starting points, we discuss how academic literature has engaged with the initiatives, both bringing them to existence through systematic promotion, and offering critical assessments of their content. In the closing section, we discuss ways forward in framing Data and AI for Good initiatives as part of the wider efforts to understand the relationship between computational technology, politics and society.

Understanding the structure of Data and AI for Good initiatives

In this section, we trace key steps in the emergence of Data and AI for Good initiatives, showing how closely they are linked to each other and try to construct a 'field' of social good. This analysis goes beyond discussions on the origins of the initiatives by linking them together and treating them as part of a single continuum. Such network building is well known in social studies of science (e.g. Latour, 1987), and here we identify the key actors, events and stages of the process for Data and AI for Good initiatives.

Table 1. Summary of Data and AI for Good definitions, proponents, and key milestones.

Initiative or brand	Definition and proponents	Key dates
Data for Good	'Data science and use that involves an ethical approach and leads to positive action, often in the service of humanity. Can be aligned with the UN's Sustainable Development Goals' (DataKind UK, n.d.) 'The Data for Good Exchange is part of a long Bloomberg tradition of advocacy for using data science and human capital to solve problems at the core of society' (Bloomberg, n.d.) 'We empower partners with privacy-preserving data that strengthens communities and advances social issues' (Facebook, n.d.)	2011 launch of Data Without Borders (the precursor to DataKind) 2014 First Bloomberg Data for Good Exchange took place 2018 First UK Data4Good conference
Data Science for Good	'Our mission is to foster the use of data science for positive social impact. We train and support a new generation of data scientists, provide direct support to non-profits, and develop tools that ensure data science and artificial intelligence are used to positively and equitably benefit people all over the world' (Data Science for Social Good, n.d.)	2013 First Data Science for Social Good summer fellowship at University of Chicago 2014 Data Science for Social Good adopted as the theme of ACM's annual KDD conference
AI for Good	'The goal of AI for Good is to identify practical applications of AI to advance the United Nations Sustainable Development Goals and scale those solutions for global impact' (International Telecommunication Union, n.d.) 'Through research, engineering, and initiatives to build the AI ecosystem, we're working to use AI to address societal challenges' (Google, n.d.) 'The AI for Social Good workshop will focus on social problems for which artificial intelligence has the potential to offer meaningful solutions. The problems we chose to focus on are inspired by the United Nations Sustainable Development Goals (SDGs)' (NeurIPS conference proceedings)	2016 AI for Good Foundation launched 2017 First AI for Good Global Summit by ITU 2018 Google launches AI for Social Good program to consolidate ongoing work 2018 NeurIPS conference starts hosting AI for Social Good workshops

Table 1 gives an overview of some of the key definitions, proponents and dates of Data and AI for Good initiatives. It serves to demonstrate the continuity and convergence of missions across three prominent usages of Data and AI for Good. From a chronological perspective, Data for Good was the first initiative to emerge, starting to take shape in 2011 with the launch of 'Data Without Borders', which was at first a single hackathon-style event, but quickly gained traction, adopting the Data for Good slogan and organising its work under the organisation 'DataKind'. It is from this professional and civic-hacking type volunteering that the idea first emerges to explicitly frame the use of computational techniques for social purposes as 'for Good'. This communicative strategy was quickly adopted by universities and researchers, who started to frame similar work as 'Data Science for Social Good', first in a fellowship program by University of Chicago in 2013, and more prominently as the designated theme of the ACM's annual Conference on Knowledge Discovery and Data Mining in 2014. In several histories of Data and AI for Good initiatives, the 2014 conference marks the first time the moniker comes to wider attention and is therefore often taken as the moment of origin for the movement. 'AI for Good' is the last version to emerge, with a foundation by the same name launched in 2016 and the UN adopting in 2017 'AI for Good' as their preferred concept of referring to computational methods serving Sustainable Development Goals.

From the initial starting points, the use of the Data and AI for Good concept started to spread, solidifying itself in public and academic use as a *specific field of using computational methods* (for discussion, see Berendt, 2019; Moore, 2019;). The publication of systematic literature reviews on the *field* of AI for Good initiatives (Cowls et al., 2021; Shi et al., 2020) is an important milestone in how the moniker has moved from being a catchy slogan into a strategic concept that denotes a field of applications. This move is also evident in academic conferences like NeurIPS adopting it as a label for a specific field of study. At the same time, the UN definition of 'AI for Good' as a field delineated by sustainable development goals (SDGs) has become subject to programmatic promotion in academic journals (Floridi et al., 2020; Tomašev et al., 2020).

To critically dissect the content of the initiatives, we will break down each 'for Good' moniker by its predicate (what kind of good is pursued) and subject (what is meant as the intervention), analysing both against differences in who is pursuing the initiatives. We start our discussion from what 'for Good' stands for in the initiatives.

The fact that 'for Good' was left intentionally loose in the early days of the initiatives and still lacks precise definitions has been noted by professionals, scientific participants and critical commentators. Lack of precision has been explicitly tackled by several academic commentators who propose that the idea of goodness in the initiatives might be united only by a focus that is not exclusively profit-driven (e.g. Berendt, 2019; Green, 2019; Moore, 2019) The difficulty of defining what goodness stands for feeds into a tendency to frame Data and AI for Good both as an innovative novelty and an established agenda. By insisting that Data and AI for Good is a field of applications, proponents of the initiatives can claim that *any* project relating to a particular version of social good as defined by the proponents is part of the shared program. In research literature, Shi et al. (2020) assign all AI applications in specific domains as AI for Good. Such examples demonstrate the constant changes to the boundaries of initiatives, particularly as competing initiatives include and exclude implicit aspects of social good according to their own needs. It is thus primarily through the tools and their application that we can understand what Data and AI for Good stands for. This brings us to what kind of interventions the initiatives call for.

The subject of each 'X for Good' phrase tells us what is being used to deliver socially beneficial outcomes. In Data, Data Science and AI for Good, we see an incremental development of computational methods that use differentiated titles to maintain an aura of novelty. As shown in Table 1, organisations have proposed new variations when the public attention has turned to new topics. In the future researchers should thus be prepared to face new variations in what technologies are expected to deliver social change but should be careful of mistaking them as completely new developments.

First to emerge, Data for Good tapped into the hype around Big Data as a general phenomenon, positioning data as a source for innovative insights and interventions. At the start, we see an appeal to community-driven change in Porway's 2011 blog post that eventually led to launch of DataKind, explaining a desire 'to round up data folk who want to do something meaningful with some of their spare time and match them up with non-profits / small companies who need data services' (Porway, 2011). This ambiguity follows in the Bloomberg Data for Good exchanges, which had an exploratory focus on collaborations between corporate philanthropists and computational researchers. The first iteration of Data for Good, therefore, represented maximum ambiguity in both the form of good and the tools of intervention, being little else than a rhetorical strategy that put positive spin on the technological developments.

Data Science for Good is distinct from its predecessor Data for Good because it targeted an academic audience and promoted specific computational tools (Mann and Sahuguet, 2018). It is used by experts in the application of machine learning and computational data science methods, as opposed to generalist users of data (Paolotti and Tizzoni, 2018). Nevertheless, the boundary between data science and data is sometimes blurry, as in Juech's (2021) definition of 'Data for Good' in the edited book *Data Science for Social Good: Philanthropy and Social*

Impact in a Complex World, which defines 'Data for Good' through the use of 'big data' by non-profit organisations. The hallmarks of Data Science for Good can be seen in the Data Science for Social Good university projects. In the projects, students of computational disciplines work on projects defined by the public sector needs and corporate sponsors, such as homelessness, social and health care delivery, urban planning and racial justice. In combination, Data Science for Social Good emphasises the role of data scientists as agents of change who wield specific computational tools to contribute to social good.

Finally, AI for Social Good seeks to deploy advances in machine learning and AI research to prevent, mitigate or resolve human and natural world issues, without introducing or exacerbating harm and inequality (Cowls et al., 2021; Tomašev et al., 2020). What exactly is AI is is left ambiguous, as often happens in popular discourse. In their survey of technical research on AI for Good, Shi et al. (2020) found that AI could refer to a variety of computational techniques, but machine learning was the most popular technique in every field of application, in addition to healthcare being the most popular overall category and most dominated by machine learning,

The above overview shows an expanding field of initiatives and organisations that promote data, data science or AI as tools of achieving socially beneficial outcomes. Analysing the initiatives as part of a single continuum allows us to show the wider trajectory that connects them together. In the next section, we trace this history further back by showing how Data and AI for Good initiatives continue a trajectory laid out by the ICT for Development movement.

Continuity with ICT for development

To take an analytical step back from the promotional logic of Data and AI for Good initiatives, the initiatives should be understood as a new link in the chain of past and current movements that seek to solve societal issues with technology. We suggest it is imperative to take an empirical study of Data and AI for Good initiatives as an opportunity to identify their similarity and continuity with historical and contemporary parallels. To show the continuity between initiatives, we examine the connections between ICT for Development (ICT4D) and Data and AI for Good initiatives. The above-identified strategies of network-building through corporate, non-profit and academic collaborations were already pioneered in ICT4D and have been given extensive treatment in the literature. We do *not* argue that ICT4D is the *only* important parallel and encourage future research to identify other overlaps but suggest that ICT4D is a case where the overlaps are particularly clear and provide important starting point.

ICT4D emerged in the 1980s and expanded in the subsequent decades into a widely acknowledged professional and scholarly movement. This development ultimately led into ICT4D being used as a title for conferences, UN programs, corporate projects, philanthropic initiatives, research activities and academic journals (for the emergence of the academic field, see Walsham, 2017). The focus of ICT4D also changed from an early emphasis on modernisation and economic developments towards a more multifaceted notion of development embodied in the UN Millennium Development Goals. The initiatives showed how development organisations like the UN and large non-profits were keen to embrace Western private technology companies as financial sponsors and technology providers, adopting the positive narratives of technology driving social progress. This strategy continues in the way Data and AI for Good projects are promoted by the UN, large technology corporations and philanthropic donors as we have shown above.

ICT4D scholarship notes that the movement initially grew from researching information systems in developing countries, which had a focus on technological innovation and implementation. Nevertheless, whether and how ICT might be a vehicle for development has often remained elusive (e.g. Zheng et al., 2018). Indeed, lack of agreement on what is development (e.g. Avgerou, 2010; Crush, 1995; Zheng et al., 2018) and difficulties in defining ICT (Sein and Harindranath, 2004) have troubled the field almost from the beginning. These debates are analogous to the uncertainties regarding what is computational technology and what constitutes goodness in Data and AI for Good. Given these ambiguities, ICT4D has also struggled to establish a robust theory of how advancement of ICT will lead to development, despite this being fundamental to the entire enterprise (Best, 2010; Heeks, 2010; Walsham, 2017). These ambiguities, however, did not prevent ICT4D from becoming a major professional and scholarly phenomenon, which testifies to the power of professional and academic movements to further their goals despite criticism.

ICT4D and Data and AI for Good share common points of critique. ICT4D has been criticised for technological determinism and neoliberal expansion of markets, which work against the goals of solving economic and social inequalities or digital divides (e.g. Pieterse, 2010). Such criticism has also been levied against Data and AI for Good initiatives in the development context (Madianou, 2021; Magalhães and Couldry, 2021). The theoretical challenges relating to political economy and technological determinism, therefore, apply to both movements. Academic scholarship on ICT4D includes strong critical voices on these problems (De' et al., 2018; Schelenz and Pawelec, 2021), but professional practice has often remained resistant to these critiques. Similar challenges also mark the discrepancy between continuing hype and investment into AI in social contexts despite increasing criticism and lack of progress on the ground. The emergence of Data and AI for Good extends the experiences of ICT4D in how technical research to solve non-technical problems can sustain its momentum despite tenuousness in core theoretical assumptions and vehement social scientific critique.

Transitions from ICT4D to AI for good

The connection between the two movements sometimes takes the form of direct transition from one to the other. An empirical case for transition can be found in the International Telecommunications Union (ITU) work on AI for Good, which was already mentioned above. The United Nations, of which ITU is part of, was central to the emergence of ICT4D in the 1990s, when it adopted policies that framed technology as a vehicle of development and launched collaborations with leading private-sector ICT corporations. ITU became a central partner in UN work on ICT4D after coordinating the UN 2005 World Summit on the Information Society and launching the 'Partnership on Measuring ICT for Development'. In 2017 ITU was chosen to spearhead UN work on Artificial Intelligence by convening a recurring global summit on AI for Good. The initiative was co-sponsored by IBM, the philanthropic technology organisation XPRIZE and the US-based learned society Association for Computing Machinery (ACM). The new ITU AI for Good program defined 'goodness' through the UN sustainable development goals (SDGs), which also serves as the operational definition of ICT4D in UN work. ITU work on AI for good is therefore a direct extension of earlier UN work on ICT4D. The move to position SDGs as the thematic foundation of AI for Good has also been adopted by academic promoters (Cowls et al., 2021; Tomasev et al., 2020) and private corporate actors (Chui et al. 2018).

To assess the continuity between ICT4D and AI for Good in the work of the ITU we have conducted a thematic content analysis of official blog postings on the ITU AI for Good website. All blog posts between 1.1.2021 and 31.12.2022 (N = 88) were collected and subjected to a two-stage inductive coding process. Following Silverman (2020), blog posts were thematically coded by their overall topic of interest, the technology discussed and its application, the name and type of collaborators or partners mentioned in the blog and the SDGs which the blog identified as contributing toward. The first round of coding identified 28 thematic topics. These 28 themes were then grouped to present a more manageable set of themes and were used to code the blogs again in the second round of coding. Figure 1 shows the lead topics that were identified in the blog posts and the frequency of their appearance.

The use of AI tools and approaches in mitigating climate change and tackling environmental issues ($N = 12$), and in solving a range of 'grand challenges' that cut across the SDGs (N = 12), were the most prominent topics of blog posts during the sample period. Other prominent themes included AI education and governance in sustainability and development contexts (N = 8), and the use of AI tools and approaches in reducing health inequalities and delivering healthcare interventions, typically in the Global South (N = 10). Many of the themes shown in Figure 1 directly mirror themes of concern to ICT4D scholars and ITU ICT4D practitioners: the 2003 World Summit on the Information Society, convened and led by the ITU, was a pivotal point in the codification of ICT4D practice (Heeks, 2014; Unwin and Unwin, 2009). Central to the ITU's vision at this point was to ensure that the 'usage and deployment of ICTs should seek to create benefits in all aspects of our daily life' (ITU, 2003), with explicit

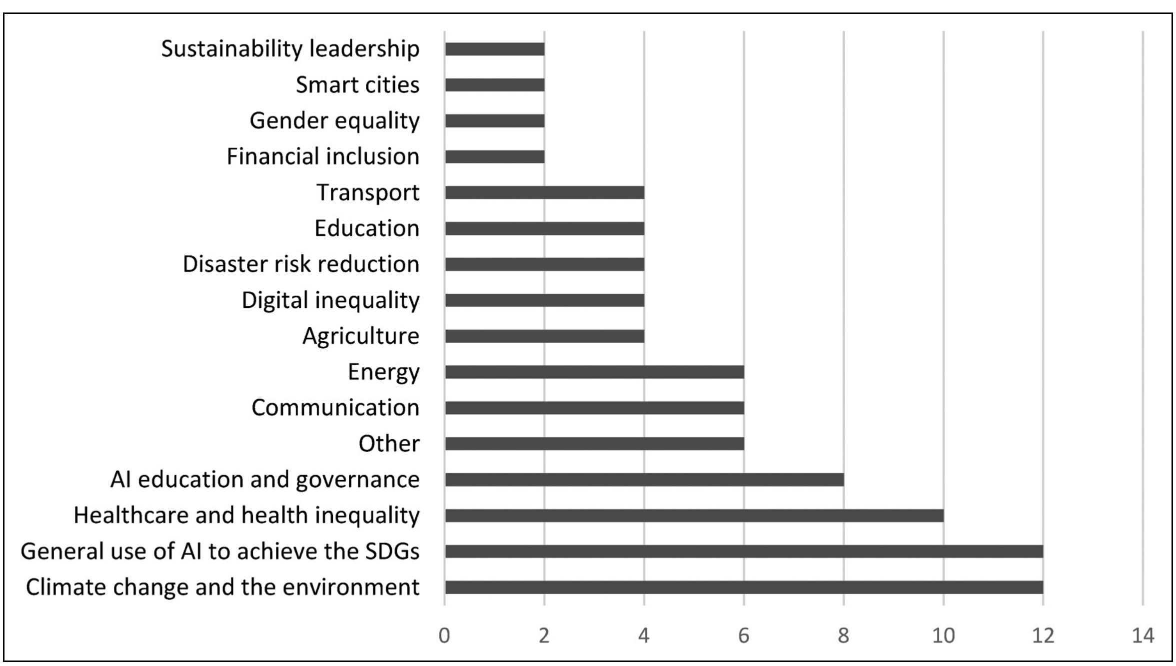

Figure 1. Themes identified in ITU AI for good blog posts between 1 January 2021 and 31 December 2022 ($N = 88$).

goals ranging from government services, employment and health to business, disaster prevention and education. An almost identical set of themes, although organised differently, is identified by Walsham (2017) as the key themes in ICT4D research. It is therefore evident that it is the same challenges which have been the focus of both ICT4D scholars and practitioners, and proponents of AI for Good initiatives.

In addition to the continuation of topical focus from the UN's ICT4D work to AI for Good initiatives, the blogs identify a long-running reliance on partnerships with external organisations that develop new computational technologies and an emphasis on the future potential rather than current practice to solve grand challenges. Strategic public-private sector partnerships were viewed as vital in ensuring ICTs were put to effective use in development contexts, and we see a high value placed on private sector organisations in the AI for Good space. In fact, SDG 9 (Industry, innovation and infrastructure) was the most frequently 'tagged' SDG across the ITU AI for Good blogs and 55% of all blogs featured private sector organisations as a collaborating partner (see Supplemental File 1). IBM featured as a notable partner in 14 (out of 88) blogs, being the most common private collaborator and a 'leading sponsor' (IBM, 2021) of the ITU's AI for Good initiative. Many of the blogs position IBM AI technologies and initiatives as central to the achievement of the Sustainable Development Goals. In many blog posts, it is the *potential* of AI technologies to solve grand challenges that is the primary focus of the article, rather than successful cases of AI technologies delivering their promises.

A second pertinent case of direct continuation between ICT4D and AI for Good can be found in Microsoft Corporation. If ITU AI for Good demonstrates how a major development actor has moved from one initiative to the other, the case of Microsoft offers a case of a private corporation doing the same. Between 2006 and 2017 Microsoft was one of the top contributors of papers to the annual academic ICT and Development conference (Mukerji and Chauhan, 2020) and Microsoft's involvement with UN-sponsored ICT4D efforts is well documented (e.g. Leye, 2007). In 2018, Microsoft launched the AI for Good Lab, which publishes research and collaborates with external partners to tackle issues such as inclusivity, sustainability and human rights, especially in the Global South. The approach chosen by Microsoft closely corresponds to those of other major technology corporations engaging in AI for Good (Magalhães and Couldry, 2021) and are closely aligned with the strategies adopted in ICT4D. Furthermore, critiques of tehnological reductivism, neoliberalism and market expansion have been levied against Microsoft's AI for Good projects (Magalhães and Couldry, 2021), just like they were made against Microsoft's involvement with the UN already two decades before (Leye, 2007).

The above analysis demonstrates the continuity and similarity between ICT4D and Data and AI for Good initiatives. Our main argument is that the way ICT4D expanded and sustained itself as a movement presents a strategy that AI for Good initiatives try to replicate. The success of ICT4D allows it to exercise influence on the understanding of technology in the entire development sector. Data and AI for Good initiatives use the same strategies and governance arrangements to promote computational technology. ITU AI for Good initiative and Microsoft's AI for Good Labs present direct evidence of organisations transitioning from one to the other.

Four approaches to Data and AI for Good in literature

Because Data and AI for Good initiatives have grown beyond a mere slogan, a budding literature has emerged to address them as more than just a disparate collection of technical applications. Here we discuss key arguments in this literature. We do not discuss the content of technical papers or project reports that align themselves with the initiatives. Detailed examination of this ever-increasing technical literature would be futile for our purposes because the inclusion and exclusion of any single project into the Data and AI for Good initiatives is an outcome of the initiatives themselves. The review therefore adopts the same strategy as some recent reviews of ICT4D literature, which only consider papers that identify themselves as part of ICT4D rather than any paper that considers ICT in developing countries (e.g. Walsham, 2017). While retroactive classificatory surveys like Shi et al. (2020) and Cowls et al. (2021) have value in mapping what is being done in the initiatives, they also promote Data and AI for Good as a unique field of applications.

In Table 2, we distinguish four varieties of literature that have so far emerged to address Data and AI for Good initiatives. These approaches show that researchers are taking a variety of positions towards the initiatives and that a research debate is clearly emerging on the topic. The literature was collected by searching key scientific publication

Table 2. Approaches to Data and AI for Good initiatives in existing literature.

	Grounded in the analysis of Data and AI for Good initiatives and their practices	Grounded in wider debate in social science and humanities
Prescriptive literature	1. Promotional literature	3. Debate on what 'goodness' ought to be
Descriptive literature	2. Analysis of Data and AI for Good as initiatives	4. Critical appraisals of Data and AI for Good projects

databases with all variants and abbreviations of Data and AI for Good initiatives and manually searching literature from key journals. We organise them according to whether they are prescriptive or descriptive in nature, and whether they are grounded in the analysis of Data and AI for Good itself or link them to wider theoretical debates in social sciences and humanities. This classificatory strategy tries to concentrate on the meta-level perspective of *how the literature relates itself to the emerging Data and AI for Good initiatives* rather than surveying their substantive contributions.

Promotional literature

The bulk of research literature has so far concentrated on promoting the existence and importance of Data and AI for Good as a field. After the initial 2014 ACM KDD conference that popularised Data Science for Social Good, researchers started to use the moniker as a keyword to position technical papers as part of the emerging field. These papers do not necessarily discuss Data and AI for Good as initiatives but concentrate on discussing specific applications and projects. Proponents of Data and AI for Good use the initiatives to organise otherwise disparate technical projects as a connected network. It thus follows that numerous technical and 'lessons learned' papers attach themselves to Data and AI for Good discussion so that they can benefit from being part of it.

Far more interesting from our perspective, however, are programmatic papers that are published in the years following the initial 2014 conferences to outline and celebrate the birth of a new research program. Earliest examples of academic promotion can be found in Catlett and Ghani (2015) and Chandy et al. (2017), which were the first journal articles to explicitly concentrate on outlining and promoting the existence of a new research program and therefore expanding and stabilising the frame introduced in the initial conferences. Some of the first pieces were directly associated with the ACM KDD conference and its adjoining Bloomberg Data for Good conference, or other early conferences that adopted the Data for Good title (e.g. Mann and Sahuguet, 2018; Paolotti and Tizzoni, 2018). Many of these papers were written by the technical researchers in private research laboratories that develop and promote computational technologies. The papers take for granted the celebratory framing of the early events and sought to affirm their goals and vocabulary. The promotional paper by Tomašev et al. (2020) in Nature Communications and Floridi et al. (2020; see also Cowls et al., 2021) are good examples of how an increasing number of researchers beyond the initial conferences adopted and promoted the initiatives. Here we also see that the promotional activities were not an instant success, seen in the long timespan between different papers that promote the same ideas.

Academic promotional literature often converges with and is supported by professional literature published by think tanks and consulting companies. For example, the UK-based NESTA adopted DataKind's understanding of Data for Good in their eponymous report (NESTA, 2015) and McKinsey adopted the vision of AI for Good that was gaining traction in Silicon Valley (Chui et al., 2018). Such dynamics emphasise the notable entanglement between academia and practice, particularly within promotional literature. For example, the call for greater cross-sector collaboration in AI for Social Good projects by Tomašev et al. (2020) is made by authors with affiliations to both academic institutions and leading commercial proponents of artificial intelligence applications.

The promotional literature can be compared to the early literature on ICT4D, because its goal is to stabilise the existence of Data and AI for Good as a field of study. Most of the promotional literature only comes *after* private corporations and non-profits have started to promote the initiatives, just like much of the academic ICT4D literature followed rather than preceded initial private and non-profit involvement. The promotional literature, therefore, starts from the idea that a field it promotes already exists, although it is the literature itself that is for its own part bringing the initiatives into existence.

Debate on what 'goodness' ought to be

The success of the Data and AI for Good initiatives soon prompted reflection on how to define the idea of social good that the initiatives were pursuing. This literature parallels that in ICT4D on the difficulty of defining development but departs from it with its abstract orientation. From the start, this literature has been a prescriptive discussion on what goodness ought to be. While many of the authors lament the ambiguity of what goodness might stand for, they are inclined to support Data and AI for Good initiatives and focus on elaborating on the qualifications that would form a solid foundation for them. This often places them surprisingly close to the promotional literature discussed above, but distances them from critical social science literature.

In some of the first papers to offer a critical reading of Data and AI for Good initiatives, Moore (2019) and Berendt (2019) argued that Data and AI for Good initiatives frame goodness as a domain of application. They propose that the idea of goodness in the initiatives is not founded on shared principles or political goals but defined by the domains in which computational tools are applied (such as healthcare, environment and development aid). The authors argue that such a domain-based definition exposes the initiatives to ambiguities and ethical challenges. The importance of this inductive definition cannot be emphasised enough, because it reflects the loose way that 'for Good' took form in the initial stages, as we discussed in the previous section. However, Moore's and Berendt's papers used the commentary only as a stepping stone for alternative definitions of goodness. They did not develop the argument into a more detailed discussion on the initiatives themselves or their wider ramifications. Nevertheless, their argument has wider analytical value and served as an important starting point of this paper as well.

Some authors have proposed that goodness should be constituted by following ethical standards in how projects are developed and deployed. The arguments in this approach originate from the wider discussion on information ethics and ethical data science, making Data and AI initiatives just a small area of the overall discussion. For example, the special issue on AI for Social Good in Philosophy and Technology edited by Cowls (2021) treats the subject matter only as a loose umbrella for philosophical debates. In more direct contributions to the pursuit of social good, researchers have proposed pre-defined ethical principles (Floridi et al., 2020), and practical questions to guide ethical reflection (Berendt, 2019, 2020).

Scholars have also tried to define goodness through the goals projects try to attain. Green (2019, 2020) has argued that goodness should be defined through an aspiration towards social justice and fight against oppression as a political mission. Moore (2019) offers another version of this politically motivated definition by proposing that the focus should be on augmenting human capabilities. Cowls et al. (2021) propose that AI for Good initiatives should follow the United Nations Sustainable Development Goals as a universally agreed upon guideline for what good might look like, which affirms the definition proposed by the UN-affiliated ITU in 2017. All these authors connect goodness to a wider public debate that transcends academic definitions.

Analysis of Data and AI for Good as initiatives

Many papers that examine Data and AI for Good initiatives include discussion on how these initiatives came to be, who sponsors them and what applications they promote. Nevertheless, most such descriptive arguments do not explicate the social, organisational, communicative and economic structures of initiatives. For example, Berendt's (2019) and Moore's (2019) important arguments about the domain-based definition of goodness contain discussion on how the initiatives operate, but the main thrust is in the prescription of what goodness ought to be. Furthermore, many of the critical social science papers that we will discuss in the next section include discussion of how Data and AI for Good initiatives operate, but ultimately do this to make a wider argument about contemporary datafication (e.g. Magalhães and Couldry, 2021) or to deliver an overall critique of using AI as a solution to non-technical problems (Holzmeyer, 2021). Espinoza and Aronczyk (2021) offer one of the few studies that fully concentrate on how Data and AI for Good initiatives are organised. Their study shows how projects on data and climate action have been brought in as part of the wider Data and AI for Good discussion and how they are shaped by a network of dependencies between stakeholders who promote the use of computational tools.

All the above are valuable arguments in their own strands of literature but are unsatisfactory when it comes to understanding how the initiatives themselves are positioned in a longer chain of similar attempts. The empirical contribution of our paper is to start filling this gap and provide directions for future scholarship.

Critical appraisals of Data and AI for Good projects

An increasing number of empirical and theoretical literature draws on wider social science literature to challenge the idea that Data and AI for Good initiatives lead to better societal outcomes. The literature we discuss here taps into the critical current of research that emerged in the wake of the Big Data debate in the early 2010s and stands in direct opposition to the promotional literature. Furthermore, the literature also differs from the debate on what goodness ought to be by not proposing alternative definitions of social good.

Scholars have pointed out that the goodness pursued in the Data and AI for Good initiatives is often narrowly defined through the attributes of novel computational techniques and that they ultimately serve the interests of their corporate developers (Magalhães and Couldry, 2021). When examined from a broader perspective, these authors suggest, Data and AI for Good initiatives might not only fall short from solving the social problems but perpetuate them and create new ones. The argument has been delivered in a broad critique as done by Holzmeyer (2021) and Green (2020), who connect the issue to the burgeoning literature on critical data and algorithm studies. On the other hand, Madianou (2021) and Magalhães and Couldry (2021) argue that Data and AI for Good initiatives make development aid into a privatised and corporate effort with significant colonial overtones that have a net-harmful effect. These studies are just those that explicitly address Data and AI for Good initiatives, but a similar dynamic has been detected in critical studies of ICT4D (De' et al., 2018; Schelenz and Pawelec, 2021). Because of this broad foundation, the above studies are just the tip of an iceberg in a wider critical literature on contemporary computational technologies. Indeed, critical research on ICT4D provides an important starting point for research on Data and AI for Good initiatives, especially when connected to the broader critical scholarship on computational technologies in fields such as sociology, science and technology studies and media and communications.

Ways forward for research

As discussed above, literature is already rife with prescriptive studies that either promote the broad concept of Data and AI for Good or debate the correct formulation of their normative foundations. On the other hand, few descriptive studies have tried to understand Data and AI for Good as initiatives that exercise and ongoing influence on how computational techniques are understood and deployed in social contexts. Even if the influence of the initiatives is restricted to the academic community, they have an impact on how researchers understand the relationship between computational techniques and society, and the position of individual researchers in relation to politics. In this section, we underscore theoretical themes that are relevant for future studies.

We argue that the way forward lies in analytical description and critical inquiry rather than prescriptive study. In other words, we call for researchers to step back and conduct theoretically rigorous analysis of the practices, parallels and consequences of the initiatives. Future research should explore whether new private and non-profit-driven initiatives like Data and AI for Good are similar or different from the various network-building strategies known in sociology of science (e.g. Latour, 1987). This also requires comparative research with other contemporary technopolitical movements, especially those with more movements-driven or emancipatory focus such as Data for Black Lives. The potential of this approach lies in analysing academic literature and projects *as part of the initiatives themselves* rather than as substantive conversations of their own, therefore requiring a meta-level understanding of the relationship between scholarship and the initiatives. In this paper, we have taken the first steps in this direction. Special focus should be given to the economic and discursive power of the initiatives. Given that scores of critical research into ICT4D have often been unable to change the course of techno-solutionist projects, researchers should explore how the initiatives contribute to the continued positive appeal of computational technologies despite the ever-growing evidence of repeated failures. On the other hand, there is a need to analyse how corporate and philanthropic sponsorship displaces strategies drawing from social science research and critical approaches. Lastly, in the case of Data and AI for Good initiatives failing to gain further traction or being superseded by other similar initiatives, researchers should seek to explain why some initiatives succeed and others fail. Apart from being a phenomenon by itself to be analysed, Data and AI for Good projects call for inquiry that dissects how their processes and practices interact with broader theoretical themes. Much of this research is conducted under the loose header of Critical Data Studies (Dalton et al., 2016; Kitchin and Lauriault, 2014) and its parallels in social research into algorithms, AI and digital platforms. In some cases, the existing critical ICT4D literature can be directly extended to cover Data and AI for Good initiatives, as is done by Madianou (2021). The strength of these existing discussions is the primary reason why we are not calling for a launch of a new research program with its own keywords and instead call for linking the topic with existing literature. We believe that the next step in research should be in consolidating the burgeoning literature in relation to the Data and AI for Good initiatives and their parallels.

Many opportunities can be found in employing a structural perspective on the interaction between individual cases and broader sociotechnical, economic and cultural structures. The role of large technology companies and development aid industry are just starting points in the more nuanced understanding of the cultural values and discourses that permeate Data and AI for Good initiatives. Steps towards this direction have already been taken by some of the scholars that offered the first critical appraisals of the initiatives, such as Magalhães and Couldry (2021) and Madianou (2021) who explicitly link Data and AI for Good projects to power asymmetries and shared technological imaginaries. Because this strategy treats its empirical objects as potential manifestations of wider structural themes such as capitalism, neoliberalism, colonialism or technological domination, they form only a fraction of the possible wider research discussion on these topics. The critical literature on and within ICT4D provides important parallels for such research regardless of whether the empirical site of analysis is in the Global North or the Global South, although the theoretical themes will manifest in separate ways. The political economy perspective and critical discursive perspective therefore offer strong theoretical foundations for future studies of Data and AI for Good and the broader work in Critical Data Studies (Dalton et al., 2016) Data Justice (Taylor, 2017) offers a range of critical research strategies that go beyond structuralist arguments.

New empirical studies on Data and AI for Good can also potentially modify these existing theories or help identify new dynamics of computational technology and politics. Both the promotional literature and the prescriptive analyses of normativity and ethics are trying to create new formations at the intersection of technology and society. Scholars like Green (2020) and Cowls (2021) are keenly aware of political dilemmas and try to find ways forward from the current challenges, and their work is paralleled by the wider literature on AI ethics. We argue that it would be a mistake to dismiss the ability of such works to create new practices and dynamics. It thus remains an open question whether Data and AI for Good initiatives replicate existing structural dynamics, adjust them, transcend them or create completely new structures that would require altogether novel conceptualizations.

Another broader theme needing more attention is the emergence of data science as a profession, and how Data and AI for Good initiatives display a specific idea of not only technology and politics, but also computational *researchers* and politics. Throughout Data and AI for Good initiatives, practitioners with computational backgrounds are identifying themselves as data scientists, a professional identity that was undefined until recently. Scholars have suggested that data science as a profession under construction is itself a formation that needs further empirical attention (Ribes, 2019). Studies of the data science profession are particularly important given that many Data and AI for Good initiatives define goodness as a field of applications as discussed by Berendt and Moore but place their tools outside this domain. This strategy of domain independence has become a key strategy in data science more broadly (Ribes et al., 2019). Data science professionals claim that computational techniques are relevant in *any* context, and Data and AI for Good initiatives promote the view that some of these contexts are inherently good, dignifying those who work in them. From this perspective, Data and AI for Good projects become examples of how a new profession explores, expands and legitimates its potential and limits.

A corollary of the new subject position of data scientists as agents of social change is that data scientists themselves cannot avoid questions of values and politics, which is displayed in the debates over data science and AI ethics. Increasing ethical scrutiny means that data scientists are increasingly framed as ethicists. Data and AI for Good initiatives, therefore, provide a critical area of research for how ethics and politics are negotiated in data science practice that specifically seeks to promote social good. Indeed, Tanweer (2018) has shown how data scientists shape their ethical positions when confronting political questions in their work, and Neff et al. (2017) show how tensions are negotiated in practical projects. With an increasing focus on ethics, it should then not come as a surprise that Data and AI for Good initiatives are often framed as outlets to confront ethical challenges; they can act as an outward expression of the willingness to face criticism and to show that data science can be ethical and work for good causes despite its shortcomings. Nevertheless, as Powell et al. (2022) have shown for entrepreneurial Tech for Good networks, explicit pursuit of ethical practice and normative ideals can include paradoxes and practices that directly contradict the stated goals. The growing understanding of the data science profession may thus force us to consider the normatively prescriptive studies of what goodness ought to be not as external commentary but as debate that is internal to the data science profession itself. This would make the data science discussion on ethics and normativity itself into an object of empirical inquiry regardless of whether it happens by practitioners or academics, as the two are increasingly intertwined. The pragmatic turn in analysing normative ideals and justifications (e.g. Boltanski and Thevenot 2006) provides one avenue of research into exploring such developments empirically.

It is in no way certain that the arguments in the above approaches are compatible and herein lies opportunities for rigorous empirical and theoretical debate. Study of Data and AI for Good can be furthered with the help of wider research discussion on computation and society, and examination of the initiatives can likewise further our understanding of the broader themes.

Conclusions

In this paper, we have taken a step back from Data and AI for Good initiatives to propose how to make sense of them as initiatives that unfold both inside and outside academia. The initiatives are not just rhetorical flourish but have ongoing influence on how computational technologies are thought to interact with phenomena of political importance. We examined prominent examples, highlighted their connections with ICT4D, surveyed recent literature and finally offered ways forward.

The key contribution of the paper is in positioning Data and AI for Good initiatives as connected to each other and to other technopolitical initiatives, which calls us to research them as initiatives that shape the future of how technology and politics intersect, and as manifestations of broader theoretical themes. It is no longer productive to critique individual notions of what good stands for. Neither is it beneficial to repeat the mantra that we need more research on how individual solutions behave in complex real-world contexts. In our view, there is already ample literature on both, and the way forward is in consolidating and theorising our understanding. Furthermore, we reject the idea that research on Data and AI for Good as a phenomenon ought to develop its own specialist literature, such as the literature concerning ICT4D. Instead, we need to learn from the critical aspects of ICT4D literature and integrate them into broader empirical and theoretical themes. Creating either a promotional or a critical research program that only concentrates on the newest developments would only reify the normative and technological claims of the initiatives. What is required is the decentring of the initiatives and it is for this reason that researchers must take both a practical and analytical step back.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the UK Economic and Social Research Council under Grants ES/P000622/1 and ES/P000711/1.

ORCID iDs

Ville Aula https://orcid.org/0000-0001-7666-2121
James Bowles https://orcid.org/0000-0003-1473-6329

Supplemental material

Supplemental material for this article is available online.

Note

1. In this paper we do not abbreviate Data and AI for Good to maintain a critical distance from the initiatives and avoid using catchy, promotional abbreviations that have become important to the initiatives themselves, such as AI4SG or D4G. For brevity, however, we only use the ending 'for good' rather than 'for social good' that is used in some initiatives.

References

Avgerou C (2010) Discourses on ICT and development. *Information Technologies and International Development* 6(3): 1–18.

Berendt B (2019) AI for the common good?! pitfalls, challenges, and ethics pen-testing. *Paladyn (Warsaw)* 10(1): 44–65.

Berendt B (2020) (De) constructing ethics for autonomous cars: A case study of Ethics Pen-Testing towards "AI for the Common Good." *The International Review of Information Ethics* 28(6).

Best ML (2010) Understanding our knowledge gaps: Or, do we have an ICT4D field? And do we want one? *Information Technologies & International Development* 6(Special Edition Harvard Forum II Essays): 49–52.

Bloomberg (n.d.) Data for good exchange 2014. Available at: https://www.bloomberg.com/lp/d4gx-2014/ (accessed 26 November 2021).

Boltanski L and Thevenot L (2006) *On Justification: Economies of Worth.* Princeton University Press.

Catlett C and Ghani R (2015) Big data for social good. *Big Data* 3(1): 1–2.

Chandy R, Hassan M and Mukherji P (2017) Big data for good: Insights from emerging markets. *Journal of Product Innovation Management* 34(5): 703–713.

Chui M, Harryson M, Manyika J, et al. (2018) *Notes from the AI frontier: Applying AI for social good.* McKinsey Global Institute.

Cowls J (2021) 'AI for social good': Whose good and who's good? Introduction to the special issue on artificial intelligence for social good. *Philosophy and Technology* 34: 1–5.

Cowls J, Tsamados A, Taddeo M, et al. (2021) A definition, benchmark and database of AI for social good initiatives. *Nature Machine Intelligence* 3(2): 111–115.

Crush J (1995) Introduction: Imagining development. In: Crush J (ed.) *Power of Development*. 1st ed. Florence: Taylor & Francis Group, pp. 17–38.

Dalton CM, Taylor L and Thatcher J (2016) Critical data studies: A dialog on data and space. *Big Data and Society* 3(1): 1–9.

DataKindUK. (n.d.) Glossary. Available at: https://datakind.org.uk/glossary/ (accessed 26 November 2021).

Data Science for Social Good. (n.d.) Our mission. Available at: https://www.datascienceforsocialgood.org/ (accessed 26 November 2021).

De R, Pal A, Sethi R, et al. (2018) ICT4D Research: A call for a strong critical approach. *Information Technology for Development* 24(1): 63–94.

Espinoza MI and Aronczyk M (2021) Big data for climate action or climate action for big data? *Big Data and Society* 8(1): 1–15.

Facebook (n.d.) We empower partners with privacy-preserving data that strengthens communities and advances social issues. Available at: https://dataforgood.facebook.com/ (accessed 26 November 2021).

Floridi L, Cowls J, King TC, et al. (2020) How to design AI for social good: Seven essential factors. *Science and Engineering Ethics* 26(3): 1771–1796.

Google (n.d.) AI for social good. Available at: https://ai.google/social-good/ (accessed 26 November 2021).

Green B (2019) "Good" isn't good enough. In Proceedings of the AI for Social Good workshop at NeurIPS (Vol. 17).

Green B (2020) Data science as political action: grounding data science in a politics of justice. SSRN Preprint. Green, Ben, Data Science as Political Action: Grounding Data Science in a Politics of Justice . Available at: https://ssrn.com/abstract=3658431 or http://doi.org/10.2139/ssrn.3658431 (accessed 21 July 2020).

Heeks R (2010) Do information and communication technologies (ICTs) contribute to development? *Journal of International Development* 22(5): 625–640.

Heeks R (2014) ICT4D 2016: New priorities for ICT4D policy, practice and WSIS in a post-2015 world. Development Informatics Working Paper (59).

Holzmeyer C (2021) Beyond 'AI for social good' (AI4SG): Social transformations—not tech-fixes—for health equity. *Interdisciplinary Science Reviews* 46(1–2): 94–125.

IBM (2021) IBM AI for Good sponsorship focuses on AI in service of the planet. Available at: https://www.ibm.com/blogs/journey-to-ai/2021/11/ibm-ai-for-good-sponsorship-focuses-on-ai-in-service-of-the-planet/ (accessed 15 January 2022).

International Telecommunication Union (2003) Declaration of principles. Building the information society: a global challenge in the new Millenium. World Summit on the Information Society, Geneva. Available at: https://www.itu.int/net/wsis/docs/geneva/official/dop.html (accessed 18 January 2023).

International Telecommunication Union (n.d.) AI for good global summit 2017. Available at: https://www.itu.int/en/ITU-T/AI/Pages/201706-default.aspx (accessed 26 November 2021).

Juech C (2021) Building the field of data for good. In: Lapucci M and Cattuto C (eds) *Data Science for Social Good*. Cham: Springer, pp. 41–54.

Kitchin R and Lauriault T (2014) Towards critical data studies: charting and unpacking data assemblages and their work. Programmable City Working Paper (No. 2). Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112

Latour B (1987) *Science in action: How to follow scientists and engineers through society*. Milton Keynes: Open University Press.

Leye V (2007) UNESCO, ICT corporations and the passion of ICT for development: Modernization resurrected. *Media, Culture & Society* 29(6): 972–993.

Madianou M (2021) Nonhuman humanitarianism: When'AI for good'can be harmful. *Information, Communication and Society* 24(6): 850–868.

Magalhães JC and Couldry N (2021) Giving by taking away: Big tech, data colonialism, and the reconfiguration of social good. *International Journal of Communication (Online)* 15: 343–362.

Mann G and Sahuguet A (2018) With great data comes great responsibility: How data science can be applied to solve public interest problems without losing its soul. *Journal of Technology in Human Services* 36(1): 1–7.

Moore J (2019) AI For not bad. *Frontiers in Big Data* 2(32): 1–7.

Mukerji M and Chauhan U (2020) A social network analysis of ICTD conferences (2006–2017). *Information Technology for Development* 26(4): 788–810.

Neff G, Tanweer A, Fiore-Gartland B, et al. (2017) Critique and contribute: A practice-based framework for improving critical data studies and data science. *Big Data* 5(2): 85–97.

NESTA (2015) *Data for Good – How Big and Open Data can be Used for the Common Good*. Available at: https://media.nesta.org.uk/documents/dataforgood.pdf.

Paolotti D and Tizzoni M (2018) DSAA 2018 Special session: Data science for social good. 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), 470–471. Turin, Italy: IEEE.

Pieterse JN (2010) Digital capitalism and development: The unbearable lightness of ICT4D. In: Kalantzis-Cope P and Herab-Martin K (eds.) *Emerging Digital Spaces in Contemporary Society*. London: Palgrave Macmillan, pp. 305–323.

Porway J (2011) Doing good with data – data without borders. Available at: https://www.datakind.org/blog/doing-good-with-data-data-without-borders/ (accessed 26 November 2021).

Powell A B, Ustek-Spilda F, Lehuedé S, et al. (2022) Addressing ethical gaps in 'Technology for Good. *Foregrounding care and capabilities. Big Data and Society* 9(2). https://doi.org/10.1177/20539517221113774.

Ribes D (2019) STS, meet data science, once again. *Science Technology and Human Values* 44(3): 514–539.

Ribes D, Hoffman AS, Slota SC, et al. (2019) The logic of domains. *Social Studies of Science* 49(3): 281–309.

Schelenz L and Pawelec M (2021) Information and communication technologies for development (ICT4D) critique. *Information Technology for Development* 28(1): 1–24.

Sein MK and Harindranath G (2004) Conceptualizing the ICT artifact: Toward understanding the role of ICT in national development. *The Information Society* 20(1): 15–24.

Shi ZR, Wang C and Fang F (2020) Artifcial intelligence for social good: A survey. ArXiv Preprint ArXiv:2001.01818, 1–78.

Silverman D, ed. (2020) *Qualitative research*. London: Sage.

Tanweer A (2018) *Data science of the social: How the practice is responding to ethical crisis and spreading across sectors*. Doctoral dissertation, University of Washington. ResearchWorks Archive.

Taylor L (2017) What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society* 4(2): 2053951717736335.

Tomašev N, Cornebise J, Hutter F, et al. (2020) AI For social good: Unlocking the opportunity for positive impact. *Nature Communications* 11(2468): 1–6.

Unwin PTH and Unwin T, eds. (2009) *ICT4D: Information and Communication Technology for Development*. Cambridge: Cambridge University Press.

Walsham G (2017) ICT4D research: Reflections on history and future agenda. *Information Technology for Development* 23(1): 18–41

Zheng Y, Hatakka M, Sahay S, et al. (2018) Conceptualizing development in information and communication technology for development (ICT4D). *Information Technology for Development* 24(1): 1–14.