Gender data for good? Partnerships between tech companies and humanitarian and development organizations

Cookson, T.P.; Carlitz, R.

DOI
10.1177/20539517251368242

Publication date
2025

Document Version
Final published version

Published in
Big Data & Society

License
CC BY-NC

Link to publication

Citation for published version (APA):
Cookson, T. P., & Carlitz, R. (2025). Gender data for good? Partnerships between tech companies and humanitarian and development organizations. *Big Data & Society*, *12*(3). https://doi.org/10.1177/20539517251368242

Original research article

Big Data & Society
July–September: 1–14

DOI: 10.1177/20539517251368242
journals.sagepub.com/home/bds

Gender data for good? Partnerships between tech companies and humanitarian and development organizations

Tara Patricia Cookson[1] and **Ruth Carlitz**[2]

Abstract

In 2013, the United Nations called for a "Data Revolution" to advance sustainable development. "Data for Good" initiatives that have followed bring together development and humanitarian actors with technology companies. Few studies have examined the composition of Data for Good partnerships or assessed the uptake and use of the data they generate. We help fill this gap with a case study of Meta's (then Facebook) Survey on Gender Equality at Home, which reached over half a million Facebook users in more than 200 countries. The survey was developed in partnership with international development and humanitarian organizations. Our study is uniquely informed by our involvement in this partnership: we contributed subject matter expertise to the development of the survey and advised on dissemination strategies for the resulting data, which we also analyzed in our own academic work. We complement this autoethnographic perspective with insights from scholars of partnerships for development, and a practitioner framework to understand the factors connecting data to action. We find that including multiple partners can widen the scope of a project such that it gains breadth but loses depth. In addition, while it is (somewhat) possible to quantify the impact of a Data for Good partnership in terms of data use, "goodness" can also be assessed in terms of the process of producing data. Specifically, collaborations between organizations with different interests and resources may be of significant social value, particularly when they learn from one another—even if such goodness is harder to quantify.

Keywords

Humanitarian, development, data for good, gender equality, technology, partnerships

This article is a part of special theme on Commodifying Compassion in the Digital Age. To see a full list of all articles in this special theme, please click here: https://journals.sagepub.com/page/bds/collections/commodifyingcompassion inthedigitalage

Introduction

In 2013, the United Nations High-Level Panel on the post-2015 Development Agenda called for a "Data Revolution" (UN, 2013). This call and its related initiatives toward the UN Sustainable Development Goals (SDGs) have led to a dramatic increase in the volume of development-related data (Kelley & Simmons, 2019; Sandefur & Glassman, 2015). While governments generate much of this data, the past decade has seen technology companies enter humanitarian and development sectors through so-called "Data for Good" initiatives. These

[1]Development and Global Public Policy, School of Public Policy and Global Affairs, University of British Columbia, Vancouver, British Columbia, Canada
[2]University of Amsterdam, Amsterdam, Netherlands

Corresponding author:
Tara Patricia Cookson, School of Public Policy and Global Affairs, University of British Columbia, 1855 West Mall, Vancouver, British Columbia, Canada V6T 1Z2.
Email: tara.cookson@ubc.ca

initiatives encompass a range of activities that seek to harness data, analytics, and human resources in pursuit of some social or environmental benefit.

Data for Good initiatives often bring together actors and organizations that otherwise occupy different spaces and work in pursuit of different goals—that is, those who generate or analyze data, and those who "do the good." As such, we can understand many such initiatives as cross-sector social partnerships (Selsky & Parker, 2005). The SDGs emphasize the importance of partnerships—both under the rubric of SDG 17 ("Partnerships for the Goals") and through various fora that seek to engage the public and private sectors, civil society, and academia. The SDG 17 also includes a call to "enhance availability of reliable data" (Target 17.I).

Scholarship on Data for Good has proliferated over the past decade, largely serving to promote the concept or debate the "goodness" of data being generated. A recent review (Aula & Bowles, 2023) finds, however, that few studies examine how Data for Good initiatives are *organized* (c.f. Espinoza & Aronczyk, 2021). The authors call for "theoretically rigorous analysis of the practices, parallels and consequences," including drivers of success and failure, with a view to achieving "more nuanced understanding of the cultural values and discourses that permeate Data and AI for Good initiatives" (Aula & Bowles, 2023, p. 9). We respond to this call with a case study of Meta's (then Facebook) Survey on Gender Equality at Home (SoGEH), focusing on the partnership underlying the survey's development and dissemination.

The SoGEH's roots can be traced to a side event of the 2019 UN General Assembly, where Facebook publicly committed to using its vast data infrastructure to advance progress for women and girls around the world (Cheney, 2019).[1] Facebook subsequently partnered with development and humanitarian organizations focused on gender equality and financed a set of gender data projects. One of these was the SoGEH, developed in partnership with the World Bank, CARE, UNICEF, and Ladysmith. The survey was administered to nearly half a million Facebook users in over 200 countries in two waves during the COVID-19 pandemic. It was unique in its geographic reach and ability to generate data at a time when face-to-face data collection was largely impossible. The project thus generated "small data" (Chen et al., 2016, p. 286) on a large scale by harnessing the infrastructural and social resources of a "big data" social media platform. The data were subsequently made available through an online public portal. Despite its global reach and being launched with great fanfare (by then-chief operating officer Sheryl Sandberg), a fairly small number of academic articles and stories in the media suggest somewhat limited use of the data.

Our study of the SoGEH through the lens of partnerships is primarily interested in how social dynamics bring data into being. The research is uniquely informed by our own involvement in this partnership. We contributed subject matter expertise to the development of the survey and advised on strategies for dissemination of the data. Following the partnership's conclusion, we engaged with the data as researchers, publishing an academic article and incorporating it into our classroom teaching. This autoethnographic perspective enhances our analysis, which relies on Vestergaard et al.'s (2021) assessment framework for development partnerships and a practitioner-oriented framework focused on data use (Custer & Sethi, 2017). The former helps assess the potential of a partnership for achieving a development goal, while the latter interrogates whether and how development data gets used. Taken together, these frameworks provide us with new insight into the consequences of a Data for Good partnership.

We find that including multiple partners can widen the scope of a project such that it gains breadth but loses depth. In addition, while it is (somewhat) possible to quantify the impact of a Data for Good partnership in terms of data use, "goodness" can also be assessed in terms of the process of producing of data. Specifically, collaborations between organizations with different interests and resources may be of significant social value, particularly when they learn from one another—even if such goodness is harder to quantify.

This article proceeds as follows. The next section presents the state of knowledge on Data for Good initiatives, including those focused on generating gender data, and discusses these in the context of partnerships for development. "Data and methods" section presents our empirical strategy and introduces the case study. "Evidence of impact: data use" section presents evidence of SoGEH's impact and use. "Discussion: Understanding impact" section discusses the impact and use by bringing together Vestergaard et al.'s (2021) and Custer & Sethi's (2017) frameworks, and "Conclusion" section concludes with implications for research and practice.

"Data for good," gender data, and partnerships for development

This section provides a brief overview of Data for Good and discusses the state of knowledge on data use. We then discuss the subset of Data for Good that the SoGEH aimed to generate gender data. We conclude with a discussion of partnerships for development. To date, these topics have largely been studied separately, despite common themes.

Data for good

The term "Data for Good" originated in a one-off hackathon-style event in 2011 on Data Without Borders. Professional and civic-hacking type volunteering characterized subsequent such initiatives. Universities came on board in 2013, when the University of Chicago launched its Data Science for Social Good fellowship. In 2014, Bloomberg launched its first Data for Good Exchange, bringing together participants from across academia, industry, government, and NGOs (Aula & Bowles, 2023). Data for Good initiatives have subsequently proliferated to include private sector, academic, nonprofit, and public sector initiatives.

Importantly, there is no consensus on "goodness." Approaches vary by actor, ranging from domain-based definitions, such as improving community health or environmental outcomes (Espinoza & Aronczyk, 2021), to the promotion of social justice, to ethical standards guiding project development and deployment. The activities associated with doing good are often aligned with the SDGs, with goodness and profit-making seen as mutually reinforcing (Olwig, 2021; Espinoza & Aronczyk, 2021).

This paper is concerned with Data for Good activities initiated by social media platform companies. Such initiatives proliferated during the COVID-19 pandemic, with tech companies publishing formerly private big mobility datasets and launching related efforts under the banner of fighting the spread of disease (Walsh, 2023). Yet even before 2020, Facebook was already making big datasets available for humanitarian responses through their Displacement Maps and facilitating the sharing of information about, *inter alia*, election violence and earthquakes, through crowd-sourced initiatives (Mulder et al., 2016).[2]

Scholars have studied the consequences of humanitarian data partnerships in different ways. Concerns about "technocolonialism" (Madianou, 2019) and "data extractivism" have been raised with regard to private sector actors when humanitarian values come into conflict with commercial interests in gaining access to people, places, and data (Schröder-Bergen et al., 2022; Sandvik, 2023). There is a risk that Data for Good projects "might not only fall short from solving the social problems but perpetuate them and create new ones… [making] development aid into a privatised and corporate effort with significant colonial overtones that have a net-harmful effect" (Aula & Bowles, 2023, p. 8). Yet these tensions are not easily sorted into binaries of "local" and "commercial," or good and bad actors. The participation of big tech companies in Data for Good initiatives may be an instance of philanthro-capitalism, in which charitable activities help sell products (or increase product use, with social media) (Burns, 2019). At the same time, "local" communities may find significant value in the resources that commercial interests bring to them (Schröder-Bergen et al., 2022).

Data use

Appraisals of Data for Good projects often overlook data use. This is a considerable oversight given that the ultimate "goodness" of the data generated is arguably a function of its impact on decision-making (Open Data Watch, 2018). Here, the broader literature on information, accountability, and citizen engagement is useful (e.g., Fox, 2007; Kosack & Fung, 2014). This work departs from the idea that making information available improves governance. Empirical evidence for the link between transparency and accountability is mixed at best, leading researchers to turn their attention to identifying the *conditions* through which the public release of information can lead to better outcomes (Lieberman et al., 2014).

Most pertinent to the present study, Custer and Sethi (2017) analyze a dynamic in which reams of development data around the world go unused, ending up in "data graveyards." Their framework draws attention to assumptions that data producers make about their intended data users. These assumptions have a bearing on whether the data get used as intended by development actors such as government officials, civil society organizations, development partners such as UN agencies and international financial institutions like the World Bank, and ordinary citizens. The framework (discussed in greater detail in "Data and methods" section) helps us trace the causal logic of getting from data to action, or uptake by intended users.

This framework was originally proposed for official data that informs development financing, policymaking and programming—specifically, official statistics, and administrative data produced by national line ministries and development partners. It is motivated in part by the sheer increase in *volume* of development data, as countries in the global south are increasingly motivated to track progress toward the SDGs and orient themselves to more "open government" (Carlitz & Mclellan, 2021). In this paper, we adapt Custer and Sethi's (2017) framework for understanding the reasons why development data get used or goes unused to analyze a partnership for development (Vestergaard et al., 2021) intended to generate Data for Good. We focus on data use development and humanitarian partners, governments, journalists, researchers, and ordinary citizens.

Gender data

Gender data are one of the first forms of Data for Good to gain prominence. In July 2012, Gallup hosted a conference on "Evidence and Impact: Closing the Gender Data Gap" where then Secretary of State Hillary Clinton, World Bank President Jim Kim, and other leaders spoke on the need for relevant data to support policies to better the lives of women and girls (Demirgüç-Kunt & Randall, 2012). While these calls originally focused on the need for more sex-disaggregated data in official statistics (Fuentes & Cookson, 2020) the movement has subsequently incorporated a broader range of data and evidence such as qualitative sources and technology-facilitated citizen-generated data (e.g., Suárez Val et al., 2023; Cookson & Fuentes, 2025).[3]

A growing gender data for development industry have also seen the entrance of technology companies, including social media platform companies and those in the telecommunications sector (e.g., Jeffrie, 2023). The potential pitfalls of "big gender data" in development and humanitarian work are primarily discussed in terms of algorithmic bias and data misuse. Gender data derived from mobile phone use or social media, for example, may reproduce the deep-seated beliefs and structural realities of the world in which they are produced when women are less likely to own and access these (Vaitla et al., 2020). Others show how big data can be used in ways that violate women's autonomy and human rights (Gurumurthy & Chami, 2022). Calls for "data feminism" seek to address such issues through justice-informed design, analysis and visualization practices (D'Ignazio & Klein, 2020).

Partnerships for development

Efforts to generate gender data, and other forms of Data for Good, typically manifest as cross-sector social partnerships, which management scholars define as "cross-sector projects formed explicitly to address social issues and causes that actively engage the partners on an ongoing basis" (Selsky & Parker, 2005, p. 850). A more recent literature discusses these in the context of "partnerships for development." Vestergaard et al. (2021) map out a framework for assessing the impact potential of these partnerships.

We are interested in the specific set of partnerships for development that generate Data for Good. Writing about cross-sector partnerships between nonprofits and businesses, Henriksen (2024) argues that the nonprofit humanitarian actors end up needing to align their interests and the interests of the populations they serve with those of the commercial actors and their "social impact" strategies. Nonprofits experience "adjustment of needs to match solutions" given power imbalances in the relationship, and over time, the incorporation of business values into humanitarian action risks shifting accountability measures away from needy populations and toward business interests (Henriksen, 2024, p. 72; see also Olwig, 2021). The partnership frameworks discussed here stop short of assessing impact (in our context, data use), for which we employ Custer and Sethi's (2017) framework.

Data and methods

This section introduces the SoGEH, the object of our case study. We also describe our role in developing, disseminating, and analyzing the data produced by the survey. This autoethnographic perspective informs our application of an assessment framework that integrates the assessment of partnerships for development with a framework for understanding data use. This integrated perspective reflects our core interest in understanding the social dynamics that bring data into being.

The survey on gender equality at home

The SoGEH was one output of Project17, a Facebook initiative that took a partnerships approach to the SDGs, and for which gender equality was the first area of focus (Levine, 2020). We understand the SoGEH as a "typical case" of a Data for Good partnership, which we can use to illustrate general trends and elucidate mechanisms (Van Evera, 1997). Meta is one of the "Big Five" tech platform companies; hence, the insights we generate have broader implications.

To develop the SoGEH, Meta worked with the World Bank, CARE, UNICEF, UN Women, EqualMeasures 2030 (a data-focused civil society coalition), and the research consulting organization Ladysmith, with which both papers' authors are/were affiliated. Ladysmith had been working with Project17 prior to the SoGEH and coauthor of this paper, Cookson, was in the room at a UN General Assembly side event when Sheryl Sandberg announced Facebook's five-year commitment to using its data to advance progress on SDG5 (Cheney, 2019). Ladysmith participated in refining the conceptualization of the SoGEH, including topics it could address and development of survey questions, and subsequently analyzed the survey results, wrote public-facing reports, disseminated findings through webinars, media engagements, and an opinion editorial, and worked with a design firm contracted by Project17 to develop an online data dashboard (www.equalityathome.org). The other actors in the partnership provided inputs into the survey design, including topics covered and survey questions.

The SoGEH was administered via the Facebook platform, where users were invited to participate through advertisements. Questions touched on demographics, gender norms, decision-making, and time and resource allocation across household members. The first round was administered at the height of the COVID-19 pandemic in July 2020; it captured 461,748 respondents in 208 countries, territories, and islands. Facebook launched the first-round results with a high-level report and webinar, which included Sheryl Sandberg's presentation of the survey findings. A subsequent round was fielded in August 2021, capturing responses from 96,000 Facebook users in 200 countries, islands, and territories. The resulting country-level data are publicly available through a data portal that the authors of this paper helped design and promote. The microdata are available to researchers and practitioners upon request.[4]

Positionality and autoethnographic perspective

This study is uniquely informed by the authors' positionality and autoethnographic perspective. Autoethnography is an empirical research approach that harnesses personal experience to generate understanding of broader social and cultural dynamics (Ellis et al., 2011).[5] It acknowledges that researchers bring their own experiences, assumptions and biases to the research, in turn shaping knowledge production. This approach offers the possibility of examining social and cultural phenomena from a range of standpoints (Ellis et al., 2011) and ranges from deeply personal and narrative accounts of solo activity (Holdsworth, 2022) to reflections on the dynamics of established organizations (Sambrook & Hermann, 2018) and emergent, action-oriented civic projects (Knox, 2024). Yet while autoethnography is an established social science method, its limitations are routinely acknowledged including by its very practitioners, particularly regarding challenges of standardization and rigor (Le Roux, 2017).

In this paper, we reflect on the SoGEH through our standpoint as scholars with training in critical theory and quantitative data analysis, and as practitioners who have worked with public and private development institutions. Coauthor 1 (Cookson) began working with Project17 in October 2019 as director of Ladysmith, which Facebook engaged to advise on their gender data efforts. During that engagement, she assumed a tenure-track position in Gender, Development and Global Public Policy. Coauthor 2 (Carlitz), then a tenure-track assistant professor of Political Science, was engaged as a consultant by Ladysmith in July 2020 to help develop a report and regional briefs based on an analysis of the data produced by the first wave of the survey, under leadership of and in collaboration with the Ladysmith team. She was subsequently engaged in November 2020 to assist with developing a data visualization platform for the SoGEH. We thus bring insider reflections on a Data for Good cross-sector partnership with certain constraints: as official partners to Facebook, we signed a nondisclosure agreement that we honor in this paper.

After our formal engagements with Meta ended, we turned to the SoGEH data as independent academic researchers. We (and additional coauthors) wrote a paper on "time poverty" that we submitted to *Gender & Society* that was ultimately rejected for being "very descriptive." We subsequently wrote a new paper on social norms that were published in February 2024 in [Journal name redacted] after peer review spanning 16 months. Approaching the SoGEH data as scholars rather than hired consultants exposed us to challenges related to its utility—which speak to overall shortcomings with the partnership.

Beyond our practical experience, we engage in this study as feminist researchers who think critically about institutional power. We are well-acquainted with literature on the cooptation of feminist movements by powerful state and corporate actors and have thus been attentive in our analysis to this possibility. At the same time, our direct participation in the SoGEH (and similar initiatives) makes us attentive to the complexity and trade-offs involved in humanitarian partnerships and introduces nuance to our analysis. For example, our intimate awareness of the power dynamics within and between involved organizations, earned through years of practitioner engagement in the field, helps us avoid a simplistic framing of good (humanitarian) actors versus bad (technology) actors.

Analytical strategy

We draw on five additional metrics of the partnership's impact, which structure "Evidence of impact: data use" section: 1) instances of SoGEH's use in peer-reviewed academic publications and in grey literature; 2) mentions of the SoGEH in the media; 3) engagement with the SoGEH on social media (Facebook, Twitter, LinkedIn and Instagram); 4) use in development and/or humanitarian outputs; and 5) use cases identified on Meta's website.[6] To find academic articles, we conducted a Google Scholar search for "Survey on Gender Equality at Home," and used the same phrase to search JSTOR, Project Muse, ScienceDirect, and Scopus. To search for mentions of the SoGEH in online media, we conducted a similar Google search. We subsequently used the built-in search function for Facebook, Twitter, LinkedIn and Instagram to search for "Survey on Gender Equality at Home" and "SoGEH." We conducted these searches between February and September 2024.

We study the SoGEH initiative through two established frameworks, which we integrate to generate new insights. First, we apply Vestergaard et al.'s (2021) Partnerships for Development Framework, which seeks to assess a partnership's "impact potential" for achieving development goals (see Figure 1).[7] At its core is an "impact value chain" that connects the partnership's focus issue and ultimate output, in terms of measurable results and deliverables. This framework highlights the *collaborative potential* of a given partnership or "the potential of the partnership to create internal value, which is critical to [its] long-term sustainability" (p. 9). The authors argue that collaborative potential is a function of *linked interests* (whether the actors engaged in the partnership agree on the societal issue being addressed, and how well their missions are aligned, particularly in terms of their goals and intended beneficiaries); *resource complementarity* (in terms of the resources and capabilities of each partner)*; synergy* (in terms of the "throughput," or "actual dynamics, execution, and implementation process of the partnership" (p. 7)); and *mutual transformation* (the outputs produced, in terms of measurable results and deliverables). We find the concepts of linked interests and resource complementarity particularly relevant.

The Partnerships for Development framework allows for an assessment of impact *potential* but stops short of outlining how researchers might assess the actual impact of a given partnership. Moreover, this framework does not speak to the particularities of a partnership intended to generate Data for Good. While there may be varied understandings of impact in this context, we consider it in terms of data use by development actors. To understand who the intended

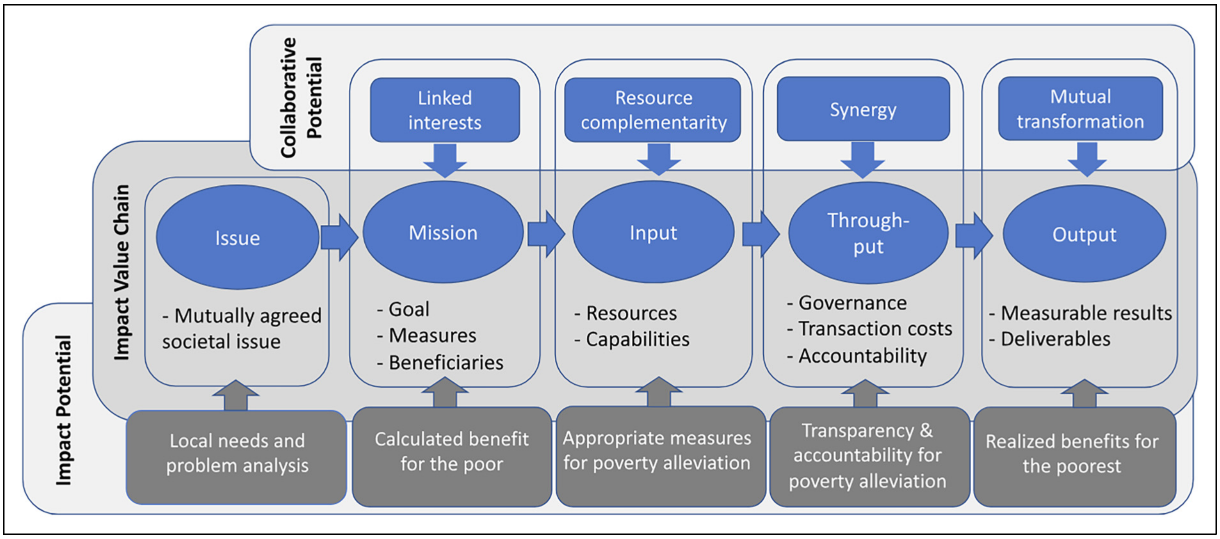

Figure 1. Vestergaard et al.'s assessment framework for partnerships for development.

users were for the SoGEH (something not explicitly stated in project documents), we note that Meta's broader Data for Good endeavors aim to, "empower partners with privacy-preserving data that strengthens communities and advances social issues." These partners include "hundreds of organizations across every continent, including universities, non-profit organizations, and international institutions." (Meta, n.d.b).

We assess SoGEH data use through Custer and Sethi's 4C framework, which considers *(1) content,* "information must be timely and salient to end users"; *(2) channel, "*easy to access and use"; and *(3) choice: "*accompanied by credible outlets for people to act upon it" (Custer & Sethi, 2017, p.14). We do not discuss *(4) consequence* (that the actions data users take "must be sufficient to change how policies are designed or programs delivered") because identifying changes to organizational strategy or behavior is beyond the scope of our study. A key component of the 4C framework is the consideration of ideal users, given that data funders and producers too often operate with "vague, and arguably naïve, archetypes of their ideal users" (ibid., p. 3).

Evidence of impact: Data use

In what follows, we present evidence of SoGEH data use through the channels identified in "Analytical strategy" section (see Appendix 1 for the full list).

Academic research

We find limited evidence of SoGEH data use in academic research. Of the 14 research articles we identified, only three use the microdata (two of which are authored by members of partnership organizations). Regarding microdata, one of the research articles (Batu & Seo, 2022) notes, "In spite of the possibility of using a rich micro-level data for our analysis, due to reasons of privacy, the publicly available data from Facebook is only at the country level. This limitation in the data forced us to estimate [country-level equations]." As we discuss below, this suggests a lack of clarity in the access regime, since Meta's publicly stated aims are to make microdata available upon request for scholars and practitioners.

Media use

Media coverage of the SoGEH was also limited. We identified 14 articles, published online in a range of outlets (see Appendix 1). We only found four mentions of the survey data on across media source platforms YouTube and Vimeo (2844 total views). The most popular of these (1700 views) was a YouTube video entitled, "How the Pandemic affected livelihood of Women in India | Barkha Dutt interview" posted by *Mojo Story*, an account that has 1.47 million subscribers. This interview was originally aired live, and the views are of the resulting post. The next most-viewed YouTube appearance was the "Survey on Gender Equality at Home: An Introduction," posted by Meta partner Tech Change (987 views). Other videos had far fewer views.

Social media engagement

Social media engagement with the survey was quite limited. Excluding posts from individuals directly involved in the SoGEH, we found four original posts on Facebook, which generated a total of 10 likes. We found five original posts on Twitter, which generated 16 reposts, and 38 likes. There was one LinkedIn post, from the Centre for Justice, Rights and Policy. We found no posts on Instagram.

Since release of the survey data, there are few instances of UN Women, CARE, UNICEF, World Bank, and EM2030 publicly promoting the partnership or its outputs, though as discussed below, the survey data have been used internally by some of these organizations.

Use in development and/or humanitarian outputs

A Google search for "Survey on Gender Equality at Home" captured 28 relevant hits (excluding hits on Meta's website). Of these, 10 were instances of development or humanitarian organizations linking to the SoGEH report(s), dashboard or microdata contact (i.e., SoGEH resource repository). Three were blogposts discussing the SoGEH's findings. SoGEH was cited in eight reports and two working papers written by development or humanitarian organizations. Lastly, SoGEH was cited in a UNESCO e-book that utilized survey data, and in a workshop on how to use survey results led by Gender Data Network.

A range of advocacy-oriented outputs use the data to document the pandemic's impact on women and call for policy changes. A report authored by 10 development actors, including the Gates Foundation, UN Women, the International Labour Organization and others, cites the SoGEH data as a "promising" example of integrating gender into non-traditional data analyses (McDougal et al., 2022, p. 12).

Meta's documentation of impact

Meta tracks "examples of how our tools have been used" on a page dedicated to the Impact of the Data for Good Initiative.[8] This page includes six examples that cite the SoGEH, which range from a report on "Data Innovation in Demography, Migration and Human Mobility" produced for the European Commission, to news stories that we have captured above. We note that one of these examples is the survey report produced by Meta itself (to which we contributed).

Discussion: Understanding impact

The previous section suggests that while the SoGEH's impact in terms of data use is fairly limited, there is some enduring impact particularly in grey (nonscholarly) research and advocacy-oriented publications. To account for and understand this impact, we apply adapted versions of Vestergaard et al.'s Partnerships for Development and Custer and Sethi's 4C frameworks.

Data use as conditioned by content, channel and choice

To more fully assess the impact of the partnership underlying the SoGEH, we reflect on the content, channel, and choices potential users made in relation to the SoGEH data.

Content. Custer and Sethi (2017) emphasize that prospective data users need to believe that the content of data is "fit-for-purpose," which relates to the perception of "the accuracy and consistency of the data being produced, as well as its perceived salience and interoperability" (p.38).

Timeliness is a key factor in the perceived salience of a dataset. The SoGEH's implementation and dissemination during the Covid-19 pandemic are commendable given that it was otherwise very challenging to collect gender data via conventional methods. The salience of SoGEH in this regard is reflected in its use by a range of development and humanitarian organizations to call attention to the (unintended) impacts of various pandemic policies on women and girls around the world.

Granularity is another key factor in salience, both in the context of a "gender data" revolution and for increased calls for intersectional and "inclusive" data that capture a range of different identity features related to group inequalities.[9] On this account the SoGEH dataset includes measures of gender, age and geography, which allowed for disaggregation and exploration of inequalities based on various identity markers. That said, disaggregation by certain categories was not possible as Meta withheld data for categories with too few responses so as not to violate respondents' privacy. If too few respondents from a given country/region/identity group responded to a given question, it might be possible to trace who these respondents are. This speaks to an access vs. privacy tension that working with big tech companies often entails.

A third factor in favor of the SoGEH's salience is that it provides data on a topic–social norms—for which there is mounting interest among development organizations and a significant degree of dissatisfaction with existing datasets (Cookson et al., 2023). The World Bank uses the SoGEH to illustrate additional insights that can be gleaned by moving beyond measurement of attitudes and beliefs (commonly used as proxies for norms) to examine the gaps between personal beliefs and social expectations (which the SoGEH does) (World Bank, 2022). Notably, the SoGEH revealed widespread (though not universal) beliefs that boys and girls should share household tasks equally, and that most people find it acceptable for women to use mobile phones.

We could not assess whether the SoGEH is being used to guide actual *programming* decisions but note the data do not score well on interoperability, meaning that the data it generate do not easily integrate with many of the official statistics and administrative datasets that development actors typically use. Lack of interoperability is a common barrier to data use (Custer & Sethi, 2017, p.16). This doesn't necessary inhibit use for advocacy, however, for which we have relatively more evidence of the SoGEH's use (discussed above).

As academics, the data provided value given its focus on topics that have not been covered in other surveys; it also provides representative data from a large swathe of the global population at the height of the COVID pandemic. However, it was challenging to construct and answer research questions using the data because the survey covered a broad spread of gender equality topics thinly (e.g., food security, violence, employment, access to and control over resources, pandemic experiences, boys' and girls' education, and opportunities). The few data points generated on any one topic made it hard to "dive deep." This broad-but-thin spread can be explained by the dynamics of a partnership including many actors with distinct priorities *within* the (broad) topic of gender equality. As an example, the World Bank is primarily concerned with economic development while UNICEF is mandated to protect children's interests. Accommodating a range of interests comes into tension with the quantity of questions that an online survey can accommodate and avoid respondent drop-off.

As authors of Facebook's reports on the SoGEH, we were able to conduct descriptive analysis useful for advocacy around gender equality. However, as academics, we struggled to formulate research questions that could be answered by the dataset alone *and* contribute to scholarly knowledge.

Channel. Channel influences impact in terms of public availability of data and how much technical knowledge is required to access it (Custer & Sethi, 2017). The SoGEH data was disseminated through three main channels: 1) a series of online reports; 2) an online "data dashboard" where users could explore the data (equalityathome.org); and 3) as microdata, available by submitting a written request to Meta.

The written reports appear to have been the most effective dissemination channel. Interestingly, the survey received the most coverage in India, where some small-scale commentary continued years later. Facebook's India office facilitated collaboration with the journalist Bharka Dhutt, which resulted in the survey receiving national coverage. This suggests the need for "data intermediaries" (Sawicki & Craig, 1996) who can process raw data into something digestible for a broader audience. Meta was willing to go beyond simply posting the data on a website and additionally invested in analysis and writeup of the most compelling findings by experts in the subject area. Investment in a media strategy at the (Indian) national level further amplified the impact.

The equalityathome.org data dashboard was intended to enhance data accessibility. Developing a usable dashboard for a public audience required thinking through a range of tensions, such as how much detail to include in the dashboard relative to data storage capacity, cost, and project timeline. For example, the more complex the dashboard, the more internet bandwidth it required to operate and the costlier it was to build. In the end, the dashboard collapses the full range of response options (e.g., "agree" and "strongly agree") for many questions. For awareness raising or advocacy by development actors, a general sense of "agreement" or "disagreement" might be sufficient. Yet for academics or development practitioners requiring nuance, the full range of answers are valuable.

As academics, we used the "cleaned" microdata available by request to Meta. It provided more detail than the dashboard, but even still, lack of access to the raw survey data posed a challenge. The complexity of the access regime, which was cumbersome and obtuse, was likely a limiting factor for broader and deeper engagement by academics. It involved signing nondisclosure and data license agreements and downloading bespoke software (which was complicated enough to require a 42-page "quick start guide" to explain it). We do not have any information on what the process entailed for other potential users, but the limited evidence of published academic work using the data suggests a major barrier.

Choice. Even when data are disseminated through appropriate channels, there can still be constraints with regard to the *choice* of individuals and groups to use it. Choice is related an actor's perception that the data is fit for purpose.

The issues we identified regarding SoGEH dissemination channels had important consequences for the choices we made as academics. As noted, the public-facing version of the dataset amounted to a subset of preselected questions and responses, collapsed such that users could not see the full variation in potential outcomes of interest. This stands in contrast to other data portals that generate considerably more engagement with researchers.[10] As a result, potential academic users of the data cannot conduct the type of exploratory analysis that typically precedes the choice of a given dataset to answer a research question. Rather, Meta, through its prespecification of which questions and responses to make public, also in a way prespecified the questions academics might ask and answer in their research.

This dynamic was not conducive to the deductive, positivist logic that guides most quantitative analysis. Quantitative studies start from a research question, develop a hypothesis and empirical strategy to answer it, and identify relevant data. In contrast, we had the data and then tried to craft research questions around it, which was time-consuming, frustrating, and ultimately unsuccessful. Using the report for Meta we had written, we attempted to "reverse-engineer" hypotheses that we had tested in our descriptive analysis. This did not add up to a coherent research article, as evidenced in our rejection from *Gender & Society*: "The broad focus of the paper (looking at a large number of state-level and individual-level factors) means that the paper is very descriptive and stays on the surface regarding the investigated issues…" While we are not averse to a more inductive, descriptive approach, it is very difficult to publish such work in "top" journals where reviewers expect a deductive, hypothesis-testing approach, and moreover tend to dismiss analysis based on observational (vs. experimental) data. This goes back as well to *content*: the SoGEH was only really set up to *describe* gender differences but not explain *why* they exist and persist.

The findings from our investigation into SoGEH data use suggest that the constraints we faced as academics may not have been as significant for actors who used the data for advocacy purposes. To date, more development actors and journalists have used the SoGEH data than have academics, particularly to draw attention to women's experiences during the COVID-19 pandemic and to call for policy changes and investments in gender equality. For such purposes, the descriptive data that SoGEH produced was useful: it provided a bird's eye view of gender relations around the world at a time when such a picture was otherwise difficult to obtain. For many forms of advocacy, descriptive data is "fit for purpose."

We did not find any examples of actors using the SoGEH dataset to inform specific policy or program design decisions. This is not surprising, given that evidence-based policy and program design requires data with high levels of context specificity and often the possibility for repeatability. In this sense, the requirements of policy and program designers have some overlap with academics.

Overall, it seems reasonable to suggest that decisions taken about the content and dissemination of the SoGEH have shaped its impact. The SoGEH data may be less useful for researchers, but reasonably useful for the purposes of advocacy. The decisions taken that shaped content and channel were a result of multiple actors with different interests and resources working in a cross-sector partnership.

Collaborative potential: Linked interests and resource complementarity

Vestergaard et al. (2021) argue that the collaborative potential of a given partnership is in part a function of *linked interests* and *resource complementarity*. In this section, we first outline the stated and implicit interests of the main actors in the SoGEH partnership and then reflect on the degree to which they can be considered "linked." We then proceed with a similar reflection on the resources each partner contributed.

(Linked) Interests? Vestergaard et al. (2021) understand linked interests as "partnership fit in terms of compatibility of strategic goals" (p. 3). Our analysis is somewhat hampered by the fact that the partners involved in the SoGEH did not explicitly articulate their strategic goals with respect to the partnership. However, we can at least partially infer these from organizational mandates and media reporting, and we can also reflect on our own goals. The actors involved can reasonably be sorted into two categories, recognizing the differences among and within these: "gender equality partners" (CARE, UNICEF, World Bank, UN Women, EM2030, and Ladysmith) and teams at Meta (P17 and data-focused).

The gender equality partners hold a range of different mandates, roles, and governance considerations. As a gender data consultancy, our interests in the partnership included having an opportunity to influence the gender data initiative of a powerful global corporation and securing paid work. Our prior engagements with Meta had been fruitful and we were animated by the prospect of producing a novel global dataset in partnership with such important gender equality organizations.

UNICEF and UN Women's programming and partnership decisions are shaped by UN norms, standards, and priorities. The UN has committed to a "data revolution" for sustainable development (IEAG, 2014). UNICEF, for example, has collaborated with technology companies "to garner insights on gender issues of relevance to children and women that are not easily captured in official statistics" since before 2020 (UNICEF, 2020, p. 3). The World Bank's "smart economics" approach views women's equal participation in the labor market as necessary for economic growth (Chant & Sweetman, 2012). The Bank's researchers explore various technology and social media company partnerships, such as Meta and LinkedIn with a view to better understanding economic trends (e.g., Marty & Duhaut, 2024). UNICEF, UN Women, and the World Bank are "Champions" of the 2018 *Inclusive Data Charter* to advance the availability and use of disaggregated data for sustainable development.[11] CARE has been a leader in gender data generation and analysis in humanitarian crises. Its Rapid Gender Assessment approach prompts humanitarian actors to make use of existing data from a range of sources and continuously update findings as new data sources emerge (Quay, 2019).

The gender equality partners also have financial interests. In 2016, UN Women committed to exploring "partnerships with the private sector, including for the development of big data projects" as a financial sustainability strategy (UN Women, 2016, p. 25). EM2030 is a cross-sector partnership initiative that promotes the generation, sharing and use of gender data and that is funded by technology companies (e.g., Salesforce) and tech-financed philanthropy (e.g., Bill and Melinda Gates Foundation). The World Bank and UNICEF both form part of the Development Data Partnership, facilitating access to data and training from Meta, LinkedIn, and Google, among dozens of other technology companies.[12]

Regarding Meta's interests, the organization had held roundtable consultations with human rights organizations to ultimately identify SDG 5 as an area in which it could contribute its unique data "superpowers" (Cheney, 2019). Citing the gender data gap as a barrier to SDG progress, the company put forth the goal to "increase the availability and use of gender data, which is critical to guiding the development of inclusive policies, programs and services, and for tracking progress on achieving gender equality" by using its "resources, data, and data science capacity" (Levine, 2020). The SoGEH was one of several resulting Data for Good initiatives.

Sharing data also entail a host of privacy and security concerns, of which companies like Meta are aware (Cookson et al., 2020). Yet such risks may pale in comparison with interests such as improving public perception and extending the reach and use of the company's products in new markets (Taylor, 2016; Whitehead & Collier, 2023). Indeed, insights into Meta's reputational interests were captured in a news article announcing the company's intention to embark on gender data partnerships: "As Facebook tries to recover from the backlash of several high-profile instances of its data being misused, the company is looking for ways to leverage its data for good while still preserving the privacy of its users" (Cheney, 2019).

Corporate interests may not always align with those of their employees, however. The individuals that make up technology companies also have personal desires to do work with a positive social impact (Taylor, 2016), and in recent years have exercised considerable power in governance decisions with regard to social issues such as discrimination within and outside the companies where they work (McGregor, 2018). Meta's COO at the time, Sheryl Sandberg, had long used her platform to advocate for women's empowerment in the workplace—an approach that became more complicated due to Facebook's various public scandals as the 2010s progressed (North, 2018).

Table 1 summarizes the interests in terms of the strategic goals of the different partners. While there is some overlap in interest in data for development and humanitarian action, there are other significant differences. For the gender equality partners, interests around mission, financial sustainability, and influence functioned in favor of tech-sector collaboration. However, early reporting on Meta's gender data partnership documented "wariness" among prominent human rights organizations about the tech giant entering the space (Cheney, 2019), including for a host of reasons covered in technology scholarship: extractive data practices baked into the business model (Espinoza & Aronczyk, 2021; Schroder-Bergen et al., 2022; Squire & Alozie, 2023) and concerns about privacy for marginalized groups. While the collaboration could improve Meta's reputation, it was not clear that it would be a net positive for the reputation of the gender equality partners. The collaborative potential of the SoGEH reflects broader challenges faced by cross-sector partnerships for development, where an "uneasy bedfellows" dynamic is common: private sector and not-for-profit and public actors partnering despite conflicting values and reputational considerations (Seitanidi et al., 2010).

Table 1. (Linked) interests of SoGEH partners.

Partners	Interests	Interests—details
Gender Equality Partners	Influence	Influence gender data initiative of powerful global tech company
	Financial	Secure paid work, financial sustainability
	Mission	UNICEF and UN Women influenced by broader UN commitments to gender equality and Data for Good; many partners publicly committed to inclusive data principles
Meta/P17	Reputation	Leveraging Data for Good as means to recover from backlash and improve reputation/market appeal
	Mission	Individuals (including former COO) saw partnership as means of furthering positive social impacts leveraging Meta's "superpowers"

Resource complementarity. Vestergaard et al. (2021) understand resource complementarity as "partnership fit in terms of the reciprocal utility of partner-specific resources" (p. 3). The Facebook platform enabled far greater reach than what is possible through conventional surveys. While it is possible to post surveys or polls on Meta for free, they would not reach a representative sample of users without the use of targeted advertising. For the SoGEH, Meta covered these costs and worked to ensure that a representative sample of users responded to both survey waves. Moreover, utilizing the platform was especially beneficial for reaching survey respondents at the height of the COVID-19 pandemic—in 80 languages. It would not have been possible for any of the partner organizations to conduct such a survey without a partnership of this type. Meta also invested considerable resources to develop the data dashboard through a partnership with Azavea, a social impact company that develops data analytics tools "for good."

The Meta researchers on the project also had deep technical expertise about the safety, ethical, and technical considerations of online surveys. This knowledge informed decisions about survey length and language translation. They also brought specific knowledge of the overall demographics of the online population in each region, which allowed for calibration of representative estimates. This latter resource was extremely useful for accurately communicating findings and analysis of a population that to date had not been surveyed at that scale.

Meanwhile, the gender equality partners brought subject matter expertise across the survey themes. This included the World Bank on employment and access to resources, UNICEF on gender norms and children's participation in education and household tasks, and CARE and UN Women on violence and the impacts of COVID policies.

Ladysmith contributed the resources of a gender research organization: we conducted gender analysis of the survey data, authored official reports and communications materials and participated in dissemination webinars and interviews. We also framed the data's contribution in terms of knowledge creation and nuanced some of the findings. For instance, when a significant share of respondents reported feeling unsafe at home, we emphasized that this *could* be a result of gender-based violence, but it could also have been a whole range of other factors for which we could not be certain—for example, respondents feeling unsafe because of the Covid-19 virus, or insecure neighborhoods. Finally, we were able to put findings that emerged from the data into conversation with existing research.

The gender equality partners also offered the project credibility as actors explicitly oriented toward “goodness.” For our own part at Ladysmith, this dynamic, the ethics, and the potential risk of the engagement, were discussed at length among team members. This resource (credibility) enhanced the collaborative potential of the partnership because of Meta's efforts toward improving public image (reputation, as discussed in interests above).

Table 2 suggests that, compared with (linked) interests, the resource complementarity of the different sets of partners was fairly high in terms of reciprocal utility. This speaks to the collaborative potential of the partnership, which we return to below.

Table 2. Resource complementarity of SoGEH partners.

Partners	Resources	Resources—details
Gender Equality Partners	Subject Matter Expertise	In-depth knowledge of survey topics, including employment, livelihoods, gender norms, household dynamics, domestic violence, and COVID impact
	Communication	Ability to communicate and share nuanced survey findings with broad development/humanitarian audiences
	Credibility	Public collaboration leant credibility to Meta's efforts given established reputation of these partners
Meta/P17	Access/Reach	Facebook platform enabled much greater reach than traditional surveys, especially during height of pandemic; Meta also invested considerably in dissemination of results through development of dashboard
	Technical Expertise	Meta employed both survey design experts and engineers with relevant expertise to develop, implement, and analyze resulting survey data

Beyond the data: Data for good as process, rather than as output

We used Vestergaard et al.'s (2021) Partnerships for Development Framework, which seeks to assess a partnership's potential for achieving impact on a mutually agreed societal issue (in this case, SDG 5). The measurable output of the partnership was (gender) data for good. We further integrated Custer & Sethi's 4C framework to go beyond impact *potential* to investigate the extent to which the data was used by intended development actors. Analyzed through these frameworks, the SoGEH appears to have had a limited impact.

Yet our experiences in this cross-sector partnership point to another way to consider the impact of Data for Good initiatives, albeit one that is harder to quantify. Specifically, we were able to learn firsthand about the possibilities, limitations, and perils of the unconventional datasets that partnerships with technology companies can offer. This impact is less about measurable outputs—datasets and instances of data use—and more about the *process of partnering*, though it is very much also about “goodness.” For us, this learning was pronounced in two areas.

The first regard privacy and safety considerations. Calls abound for open data, implying that unfettered access to data is always “good.” We learned through our involvement with the SoGEH that such calls require nuance. Meta employees spent considerable time explaining why and how to ensure that surveyed individuals are not identified. This learning has influenced our work on other projects. For example, on a technology-facilitated gender data project on the Colombian-Venezuela border, Ladysmith team members made the choice to share *analysis* rather than to openly share the project data (see Cookson & Fuentes, 2025). We took this decision based on our updated understanding of the potential risks and benefits with technology-facilitated data production, and on conversations with staff at technology companies who we met through the SoGEH partnership.

The second learning is that big tech companies—and development organizations—are not monolithic. They are composed of individuals with motivations that sometimes exist in tension with the values, intentions, or interests of the organization that employs them, and for that matter, the organizations they partner with. Data for Good partnerships have the potential to shift the attitudes, beliefs, and choices of the individuals who make up technology organizations and the teams within them. The collaborative process of developing survey questions, writing up findings, and editing communications materials offered teachable moments with regard to deploying a gender perspective—where we thought it mattered, why it mattered, and how it needed to show up. "Institutional activism" (Abers, 2021) is a useful frame for thinking about the potential for partnerships to shift the values and behaviors of organizations, in addition to their external impact, for example, on poverty or the SDGs.

In sum, the *process of partnering* to produce Data for Good through cross-sector collaboration is a key part of how such partnerships can (or can't) generate "goodness." Cross-sectoral partnerships between actors that do not often sit at the same table—that do not share the same total set of interests and resources—can be highly generative in improving future streams of data-driven work.

Conclusion

This article adds to emerging scholarship on Data for Good with a case study of a cross-sector partnership between a social media company and gender equality organizations. We assessed use of the resulting data, finding it fairly limited. Impact appears larger where there were efforts to a) translate the data into digestible findings (e.g., reports instead of access to the raw data) and b) use journalists as data intermediaries to put it in the public spotlight. These findings align with practitioner advice for "avoiding data graveyards" (Custer & Sethi, 2017).

We analyzed data use considering the partnership's collaborative potential, focusing on interests and resource complementarity. While resource complementarity was evident, interests were less well-aligned. Meta offered access to a global dataset at a time when gender data was extremely challenging to collect, but it was also dealing with serious reputational issues that caused wariness in the gender and development sector. Notably, diversity of interests was also present *among* the gender equality partners, whose mandates include different issues and constituencies. A trade-off existed between addressing each partner's interest and a dataset that enabled topical "deep dives," limiting its use for scholarship and context-specific, evidence-based policy making. Thus, one of our two main conclusions is that including multiple partners can widen the scope of a project such that it gains breadth but loses depth.

It is also possible that "gender equality" was too broad of a partnership goal. Vestergaard et al. (2021) suggest that "a [development] partnership must *intentionally, transparently, and accountably* be aimed at improving the lives of the poor" and that these criteria must be revisited throughout the partnership. A more narrowly defined goal, toward which data are collected to address a much more specific data gap—and meet the needs of a much more specific ideal user—may be more likely to avoid ending up in a data graveyard.

Our autoethnographic reflections surfaced an additional source of "goodness" in cross-sectoral Data for Good partnerships. While clashing interests often prevent big tech companies and gender equality advocates from sitting at the same table, the SoGEH facilitated learning opportunities for the individuals involved. This included a deeper understanding of important technical issues that bolster the safety, accuracy and usability of novel gender data projects. Thus, our second conclusion is that the *process* of producing of Data for Good through cross-sector partnerships is a key part of how they can (or can't) generate "goodness." This aspect of Data for Good partnerships, typically overlooked in the Data for Good literature, is harder to quantify than data use but should not be discounted.

Acknowledgements

The authors would like to thank Danielle Hidi for research assistance on this paper.

ORCID iDs

Tara Patricia Cookson https://orcid.org/0000-0003-1269-169X
Ruth Carlitz https://orcid.org/0000-0001-7840-2331

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Social Sciences and Humanities Research Council of Canada.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: We contributed in a remunerated capacity (as consultants) on the design and dissemination of the SoGEH, an engagement which ended in 2021. We did not receive any financial remuneration for the present study, which we conducted independently in our academic roles.

Notes

1. Throughout the paper, we refer to Facebook and Meta with the company name used at the time of the event referenced.
2. https://dataforgood.facebook.com/dfg/tools/displacement-maps
3. The United Nations defines gender data as sex-disaggregated statistics and data that reflect gender issues (UNSD 2016). This definition is used in Sustainable Development Goals tracking.
4. See: https://dataforgood.facebook.com/dfg/tools/survey-on-gender-equality-at-home
5. For more on the benefits and limitations of autoethnography, see Ellis et al., 2011.
6. We searched in English language only, reflecting that the SoGEH outputs are only available in English.

7. The authors "focus on development as poverty alleviation and emphasize the intentional activities taken to improve the lives of the poor" (p. 105446).
8. https://dataforgood.facebook.com/dfg/impact
9. There are emerging debates about the operationalization of intersectionality in data projects. Regarding "additive" approaches within data science, see Bentley et al., 2023. Regarding operationalization in a humanitarian gender data project, see Cookson & Fuentes, 2025. In feminist knowledge production generally, see Puar, 2020.
10. See: World Bank World Development Indicators, Afrobarometer Analysis tool, or Covid-19 Sex-Disaggregated Data Tracker.
11. https://www.data4sdgs.org/initiatives/inclusive-data-charter
12. https://datapartnership.org/

References

Abers RN and Tatagiba L (n.d.) *Institutional Activism: Mobilizing For Women's Health From Inside The Brazilian Bureaucracy*.

Aula V and Bowles J (2023) Stepping back from data and AI for good—current trends and ways forward. *Big Data & Society* 10(1): 205395172311739.

Batu M and Seo B (2022) Gender roles and safety of women at home in the COVID-19 era: Evidence from 101 countries. *International Review of Applied Economics* 36(5–6): 739–761.

Bentley C, Muyoya C, Vannini S, et al. (2023) Intersectional approaches to data: The importance of an articulation mindset for intersectional data science. *Big Data and Society* 10: 2.

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

Carlitz RD and Mclellan R (2021) Open data from authoritarian regimes: New opportunities. *Perspectives on Politics* 19: 1.

Chant S and Sweetman C (2012) Fixing women or fixing the world? "Smart economics", efficiency approaches, and gender equality in development. *Gender and Development* 20(3): 517–529.

Chen C, Ma J, Susilo Y et al. (2016). The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies* 68: 285–299.

Cheney C (2019, November 22) *Inside Facebook's emerging gender data efforts*. San Francisco: Devex. https://www.devex.com/news/inside-facebook-s-emerging-gender-data-efforts-96078.

Cookson TP and Fuentes L (2025) Gender data, intersectionality, and a feminist politics of "negotiated refusal". *Violence Against Women*. Online First. https://doi.org/10.1177/10778012241309362.

Cookson TP, Kuss MK and Bitterly J, et al. (2023) Social norms. *Gender and Development: A Review of Research and Practice. UN Women Discussion Paper Series* 42: 1–67.

Cookson TP, Zulver J and Fuentes L, et al. (2020) Building alliances for gender equality: How the tech community can strengthen the gender data ecosystem. San Francisco: Facebook and Ladysmith. Retrieved from: https://drive.google.com/file/d/1bsy_Gx7DHh8K54sRp74rJoWTLY47TdCU/view.

Custer S and Sethi T (2017) Avoiding Data Graveyards: Insights from Data Producers & Consumers. https://docs.aiddata.org/reports/avoiding-data-graveyards-report.html.

Demirgüç-Kunt A and Randall D (2012, July 19) *Small Steps Towards Closing The Gender Data Gap*. World Bank Blogs. https://blogs.worldbank.org/en/allaboutfinance/small-steps-towards-closing-the-gender-data-gap.

D'Ignazio C and Klein LF (2020) *Data Feminism*. Cambridge, MA: The MIT Press.

Ellis C, Adams TE and Bochner AP (2011) Autoethnography: An overview. *Forum: Qualitative Social Research* 12(1): 1–19. Retrieved from https://www.proquest.com/scholarly-journals/autoethnography-overview/docview/870465772/se-2.

Espinoza MI and Aronczyk M (2021) Big data for climate action or climate action for big data? *Big Data & Society* 8(1): 2053951720982032.

Fox J (2007) The uncertain relationship between transparency and accountability. *Development in Practice* 17(4–5): 663–671.

Fuentes L and Cookson TP (2020) Counting gender (in)equality? A feminist geographical critique of the 'gender data revolution.'. *Gender, Place & Culture* 27(6): 881–902.

Gurumurthy A and Chami N (2022) Beyond data bodies: New directions for a feminist theory of data sovereignty. *Data Governance Network Working Paper* 24: 1–17.

Henriksen SE (2024) Finding the "sweet spot": The politics of alignment in cross-sector partnerships for refugees. *Business and Society* 63(1): 145–184.

Holdsworth C (2022) Making autoethnography: Crafting intimate, social and material relations. *International Journal of Social Research Methodology* 27(1): 123–136.

IEAG. (2014) *A World that Counts: Mobilising the Data Revolution for sustainable Development*. Independent Expert Advisory Group on a Data Revolution for Sustainable Development. https://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf.

Jeffrie N (2023) *The Mobile Gender Gap Report* 2023. www.gsmaintelligence.com.

Kelley JG and Simmons BA (2019) Introduction: The power of global performance indicators. *International Organization* 73(03): 491–510.

Knox H (2024) Doing place through data: Proliferation, profiling and the perils of portrayal in local climate action. *Big Data & Society* 11: 3.

Kosack S and Fung A (2014) Does transparency improve governance? *Annual Review of Political Science* 17(1): 65–87.

Le Roux CS (2017) Exploring rigour in autoethnographic research. *International Journal of Social Research Methodology* 20(2): 195–207.

Levine M (2020, March 10) *Helping to Close the Gender Data Gap*. Meta. https://about.fb.com/news/2020/03/closing-the-gender-data-gap/.

Lieberman ES, Posner DN and Tsai LL (2014) Does information lead to more active citizenship? Evidence from an education intervention in rural Kenya. *World Development* 60: 69–83.

Madianou M (2019) Technocolonialism: Digital innovation and data practices in the humanitarian response to refugee crises. *Social Media + Society* 5(3): 205630511986314.

Marty R and Duhaut A (2024, May 10) *Beyond the "likes": How can we harness Facebook's marketing data for research?* Washington, D.C.DC: World Bank Blogs. https://blogs.worldbank.org/en/opendata/beyond-the–likes—how-can-we-harness-facebook-s-marketing-data.

McDougal L, Raj A, Yore J, et al. (2022) *Strengthening gender measures and data in the Covid-19 era: an urgent need for change*. https://www.paris21.org/sites/default/files/media/document/2023-10/COVID-19%20Gender%20Data%20%26%20Measures_Evidence%20Review%20V2.pdf.

McGregor J (2018, December 13) How tech workers are fueling a new employee activism movement. *The Washington Post*. https://www.washingtonpost.com/business/2018/12/13/how-tech-workers-are-fueling-new-employee-activism-movement/.

Meta. (n.d.) *About Data for Good*. Https://Dataforgood.Facebook.Com/Dfg/About.

Mulder F, Ferguson J, Groenewegen P, et al. (2016) Questioning big data: Crowdsourcing crisis data towards an inclusive humanitarian response. *Big Data & Society* 3(2): 2053951716662054.

North A (2018, December 17) How Facebook's scandals are hurting Sheryl Sandberg's Lean In brand. *Vox*. https://www.vox.com/2018/12/17/18119631/facebook-sheryl-sandberg-lean-in-feminism.

Olwig MF (2021) Sustainability superheroes? For-profit narratives of "doing good" in the era of the SDGs. *World Development* 142: 105427.

Open Data Watch. (2018) *The data value chain: Moving from production to impact*. https://data2x.org/wp-content/uploads/2019/08/Data_Value_Chain.pdf.

Puar JK (2020) "I would rather be a cyborg than a goddess": Becoming-intersectional in assemblage theory. In: *Feminist Theory Reader*, 5th ed. New York: Routledge, 405–415.

Quay I (2019) Rapid gender analysis and its use in crises: From zero to fifty in five years. *Gender & Development* 27(2): 221–236.

Sambrook S and Herrmann AF (2018) Organisational autoethnography: Possibilities, politics and pitfalls. *Journal of Organizational Ethnography* 7(3): 222–234.

Sandefur J and Glassman A (2015) The political economy of bad data: Evidence from African survey and administrative statistics. *The Journal of Development Studies* 51(2): 116–132.

Sandvik KB (2023) *Humanitarian extractivism: the digital transformation of aid, humanitarianism*. Manchester, UK: Manchester University Press.

Sawicki DS and Craig WJ (1996) The democratization of data: Bridging the gap for community groups. *Journal of the American Planning Association* 62(4): 512–523.

Schröder-Bergen S, Glasze G and Michel B, et al. (2022) De/colonizing OpenStreetMap? Local mappers, humanitarian and commercial actors and the changing modes of collaborative mapping. *GeoJournal* 87(6): 5051–5066.

Seitanidi MM, Koufopoulos DN and Palmer P (2010) Partnership formation for change: Indicators for transformative potential in cross sector social partnerships. *Journal of Business Ethics* 94: 139–161.

Selsky JW and Parker B (2005) Cross-sector partnerships to address social issues: Challenges to theory and practice. *Journal of Management* 31(6): 849–873.

Squire V and Alozie M (2023) Coloniality and frictions: Data-driven humanitarianism in North-eastern Nigeria and South Sudan. *Big Data & Society* 10(1): 20539517231163171.

Suárez Val H, D'Ignazio C, Acosta Romero J, et al. (2023) Data artivism and feminicide. *Big Data & Society* 10(2): 1–6.

Taylor L (2016) The ethics of big data as a public good: Which public? Whose good? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374(2083): 20160126.

UNICEF. (2020) *Every child counts: Using gender data to drive results for children*. https://data.unicef.org/resources/every-child-counts-using-gender-data-to-drive-results-for-children/.

United Nations. (2013) *A new global partnership: Eradicate poverty and transform economies through sustainable development. The Report of the High-Level Panel of Eminent Persons on the Post-2015 Development Agenda*.

UNSD. (2016) *Integrating a Gender Perspective into Statistics. United Nations Department of Economic and Social Affairs, Statistics Division*. New York: United Nations.

UN Women. (2016) *Making every woman and girl count: Flagship programme initiative*. https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/How%20We%20Work/flagship%20programmes/FPI-statistics-concept-note.pdf.

Vaitla B, Verhulst S, Bengtsson L, et al. (2020) The promise and perils of big gender data. *Nature Medicine* 26(1): 17–18.

Van Evera S (1997) *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press.

Vestergaard A, Langevang T and Morsing M, et al. (2021) Partnerships for development: Assessing the impact potential of cross-sector partnerships. *World Development* 143: 105447.

Walsh T (2023) Modeling COVID-19 with big mobility data: Surveillance and reaffirming the people in the data. *Big Data and Society* 10: 1.

Whitehead M and Collier WGA (2023) Incidental governmentality: Big tech and the hidden rationalities of government. *Digital Geography and Society* 5: 100071.

World Bank. (2022) *Reshaping Norms: A New Way Forward, South Asia Economic Focus*. https://openknowledge.worldbank.org/server/api/core/bitstreams/7b3cce1b-7c32-553b-83e5-0328c50d26d5/content.