DISCUSSION PAPER

Data Philanthropy, International Organizations and Development Policy: Ethical Issues to Consider

United Nations Development Programme

Contents

Introduction

In an increasingly datafied society, possessing big data means being able to save lives.[1] In emergency situations, big data can help organizations pinpoint the location of impacted individuals and allocate aid resources and response teams accordingly.[2] Most of this information can be gleaned from where people check their phones following disasters or which terms are used when communicating on social media.[3] However, big data can also assist in the handling of non-emergency situations. Companies' data can tell us not only whether a given policy intervention works or doesn't work, but also how it can be fixed.[4] As a result, a

As beneficiaries of "data philanthropy", international development organizations must consider the legal and ethical risks inherent in data sharing agreements, beyond privacy risks.

growing number of international development, humanitarian and public sector organizations have called for private digital data to be shared for the public good.[5] As a result, a "responsible data" or "data for good" movement has emerged, which calls on companies to share their data for philanthropic purposes.

The "data for good" movement promotes data-driven projects that can increase the efficiency of social initiatives, and has highlighted the significant role that the private sector can play by providing valuable data to further social action. Data generated via platforms such as telecom operators, satellite companies, and social media networks has the potential to provide a range of insights into development issues that can assist immeasurably with social research.[6] To this end, "data philanthropy" is a term that has been used to describe the giving of private sector data, access to it, or the production of data-driven insights for a socially beneficial purpose.[7] Reportedly coined by World Economic Forum Chief Technology Officer Brian Behlendorf at the World Economic Forum in 2011, the definition of "data philanthropy" remains unsettled.[8] However, this paper chooses to operate within the parameters of the definition developed by Professor Yafit Lev-Aretz, whose project is the first to engage in a privacy-focused legal analysis of data philanthropy.[9] Her project describes data philanthropy as a practice in which (1) privately-held data or proprietary data-driven insights, (2) are shared or given access to, (3) for the public good.[10]

---

[1] Alberto Alemanno, *Big Data for Good: Unlocking Privately-Held Data to the Benefit of the Many,* 9 EUR. J. OF RISK & REG. 183, 183 (2018).

[2] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1493 (2019).

[3] Annika Richterich, THE BIG DATA AGENDA: DATA ETHICS AND CRITICAL DATA STUDIES (London, University of Westminster Press, 2018).

[4] Alberto Alemanno, *Big Data for Good: Unlocking Privately-Held Data to the Benefit of the Many,* 9 EUR. J. OF RISK & REG. 183, 184 (2018).

[5] *Id*, 186.

[6] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1493 (2019).

[7] *Id*.

[8] Presently, it includes variations such as (1) the "donation of privately-held commercial data towards beneficial causes": Jane Wu, *Big Data Philanthropy: The Social Impact of Donating Data,* LINKEDIN: PULSE (Jul. 1, 2015), https://www.linkedin.com/pulse/data-philanthropy-social-impact-donating-june-wu/; (2) a "partnership in which private sector companies share data for public benefit": Andreas Pawelke & Anoush Rima Tatevossian, *Data Philanthropy: Where Are We Now?* UNITED NATIONS GLOBAL PULSE: BLOG (May 8, 2013), https://www.unglobalpulse.org/data-philanthropy-where-are-we-now; and (3) "companies sharing proprietary datasets for social good": Patrick Meier, *Big Data for Humanitarian Response*, IREVOLUTIONS (June 4, 2012).

[9] *See*, Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491 (2019).

[10] *Id,* 1498.

Organizations in both the public and private sector increasingly view data sharing as key to the success of data-driven innovation. Thus, reducing the complexities involved in data sharing is seen as crucial in enabling further collaboration.[11] However, scholars from a range of disciplines have identified a need for legal guidance on various aspects of the data philanthropy practice, especially insofar as privacy risks are concerned.[12] While guidance concerning privacy-related legal risks has been provided,[13] this paper seeks to highlight the normative risks that should be considered before entering into certain data philanthropy arrangements. In order to do this, the paper first unpacks the notion of data philanthropy and evaluates some of the key assumptions underlying the practice. The paper then discusses the normative risks inherent in sharing arrangements, including the value of the data itself, the sustainability of data philanthropy arrangements, and the privacy harms that are not presently addressed by legal regimes. As regards privacy specifically, the paper examines issues such as the risk of re-identification, the concept of collective privacy and the need for informed consent. Finally, this paper seeks to provide an initial discussion of two infrastructural modalities that address some, but not all, of the issues discussed in this paper, including the use of the United Nations (UN) Global Platform governed by the UN Global Working Group on Big Data for Official Statistics,[14] and the Safe Sharing Site proposal.

In order to provide a unique contribution to the literature, this paper narrows its focus in three ways. First, the paper concentrates particularly on the risks inherent in the provision of *personal* data, such as data collected by social media companies. This is to be contrasted with, for example, the provision of spatial mapping data. Second, it discusses the unique ethical risks involved when data is aimed at furthering *policy* development, as distinct from ad hoc emergency or disaster-response arrangements. Third, it frames its discussion of these risks from the perspective of international organizations, NGOs or civil society groups or, in other words, the *receivers* of the data. The scope of this paper is deliberately limited to what international organizations, as recipients of data philanthropy, need to consider before entering into such arrangements.[15] In so doing, the paper seeks to assist such organizations in making informed decisions about entering into data philanthropy arrangements, and does not provide a discussion of the legal or other risks that such arrangements pose to private actors providing data.

---

[11] Lisa M. Austin & David Lie, *Safe Sharing Sites,* 94 N.Y.U. LAW. REV. 581, 585 (2019).

[12] *See, e.g.* Robert Kirkpatrick, *Big Data for Development,* 1 BIG DATA, 3 (2013); Jeffrey P. Kahn, Effy Vayena & Anna C. Mastoianni, *Opinion: Learning as We Go: Lessons from the Publication of Facebook's Social-Computing Research*, 111 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES (2014); Nir Kshetri, *The Emerging Role of Big Data in Key Development Issues: Opportunities, Challenges and Concerns*, 1 BIG DATA & SOCIETY (2014).

[13] *See* Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491 (2019).

[14] *See* UN Global Working Group on Big Data for Official Statistics, *United Nations Global Platform: Data for the world* (2019), http://publications.officialstatistics.org/assets/pdf/UNGlobalPlatform_Brochure.pdf.

[15] As a separate matter, it would be useful to conduct research on broader uses of data in, for example, crisis contexts, particularly considering the Do No Harm principles, and what guidance the United Nations system could provide in this regard.

PART ONE: DATA PHILANTHROPY EXPLORED

Unpacking Data Philanthropy

The notion of "data for good" has its roots in the open data movement, which has advocated for the release of governmental data in downloadable, useable and distributable formats.[16] However, given that private actors harbor the vast majority of this data, the open data movement has expanded its target to also include private sector entities.[17] Lee-Aretz outlines a variety of different ways that organizations can engage in data philanthropy, including: (1) *Data Analytics Services* such as DataKind, which pairs high impact organizations with leading data scientists, and advocates meaningful utilization of data to solve humanitarian issues around poverty, health, human rights, education and the environment;[18] (2) *Data Storage and Data-Based Utilities* such as Microsoft Corporation's Microsoft Cloud Services which offers nonprofits and university researchers useful tools for the collection, storage and management of data, but does not involve the giving of actual data; (3) *Monetary Donations for Data Science Education and Development*; and (4) *Individual Data Sharing* which can involve (i) giving personal data about the sharing individual, such as the Personal Genome Project, which lets participants share their genome sequence and health data for the use of researchers,[19] and (ii) citizen scientists, being amateur individuals who help professional scientists speed up discoveries and innovation.[20]

> *Stakeholders in data philanthropys, problem holders, data holders and skill holders, are driven by different interests. The assumption that relationships between these stakeholders serve the public good must be questioned.*

Further, data philanthropy involves various stakeholders, each guided by individual interests. Mikel Niño and others provide a useful taxonomy to explain the different roles stakeholders can play, distinguishing between (1) problem holders, (2) data holders and (3) skill holders.[21] *Problem holders* are the individuals or institutions closest to the target population affected by the social issue to which the data relates. They may include representatives of public administration, such as governmental agencies and non-profit organizations, who work to identify the social problem, needs of affected groups, and potential solutions. *Data holders* include the individuals or institutions that hold the data, whose interest in it may arise because they collected the data themselves or acquired intellectual property or contractual rights in it. Relevantly, from a legal perspective, this category

---

[16] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1501 (2019); Tim Berners-Lee, *The Year Open Data Went Worldwide*, TED (Feb. 2010), https://www.ted.com/talks/tim_berners_lee_the_year_open_data_went_worldwide.

[17] Rajesh Chandy, Magda Hassan & Prokriti Mukherji, *Big Data for Good: Insights from Emerging Markets*, 34 J. PROD. & INNOVATION MGMT. 703 (2017); Frederika Welle Donker, Bastiaan van Loenen, & Arnold K. Bregt, *Open Data and Beyond*, 5 INT'L J. GEO-INFO. 48 (2016) (arguing that private organizations that are mandated to perform a public task and generate data in the process should not be exempted from open government data policies); Beth Simone Noveck, *Data Collaboratives: Sharing Public Data in Private Hands for Social Good*, FORBES (Sept. 24, 2015, 9:30 AM), https://www.forbes.com/sites/bethsimonenoveck/2015/09/24/private-data-sharing-for-public-good/#397107b351cd.

[18] *See* Jake Porway, *Using Collaboration to Harness Big Data for Social Good,* STAN. SOC. INNOV. REV. (Jun. 14, 2017).

[19] THE PERSONAL GENOME PROJECT, http://personalgenomes.org/ (last visited Mar. 12, 2020). *See also,* AMERICAN GUT, http://americangut.org/ (last visited Mar. 12, 2020).

[20] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1501-1503 (2019).

[21] Mikel Niño et. al., *Data Projects for "Social Good": Challenges and Opportunities*, 11 INT'L J. HUMAN & SOC. SCI. 1073 (2017), http://www.waset.org/publications/10006923.

---

does *not* include the data subjects themselves, i.e. those people to whom the data personally relates.[22] Finally, *skill holders* are those possessing the technical expertise to analyze and extract insights from the data and translate the results to the problem holders for the purpose of developing policy solutions. In the social media context, companies oftentimes operate as skill holders as well as data holders. For example, in 2016, Facebook analyzed its own data extract from the posts of Brazilian users regarding the Zika virus.[23] It then passed on these insights to UNICEF for the purposes of developing an ad campaign raising awareness about the virus.[24]

What is "Good"?

Having outlined the more structural aspects of what makes a data philanthropy arrangement, it is useful to question its key normative assumption: that these relationships function to serve the public good. The literature is currently lacking a comprehensive description of "public good".[25] Rather, scholars have pointed to discrete and ad hoc examples of collaborations that are widely accepted as being in the public interest. However, as Lee-Aretz posits, this approach is flawed. It results in a self-defining notion of "good" that is largely left to the various stakeholders, with their own interests and incentives, to describe.[26] This approach is particularly problematic considering two factors: (1) the non-rivalry nature of data and (2) the reputational gains it offers for the private sector. First, the mere notion that the giving of data is philanthropic is uncomfortable. Traditional corporate philanthropy involves a cost for the corporation (such as the giving of money) and a practical limit on the amount that can be given without harmfully interfering with the corporation's business. However, in the absence of corporations also acting as skill holders, data is offered for reuse, meaning that there is no cost to the corporation and no practical limit on the amount that can be given. In other words, by reusing data, the corporation will never run out of it. Second, data philanthropy offers various reputational benefits for corporations, which for social media companies in the present political climate may be extremely valuable.[27] Thus, in circumstances where the giving of data is done at no cost and provides data holders with a number of incentives, the ability of corporations to self-define what is "good" must be resisted. Accordingly, the data philanthropy practice calls for a comprehensive definition of social good and concrete guidelines to instruct collaborations, such that controversial characterizations of social good are thoroughly scrutinized.[28]

---

[22] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1509 (2019). US law does not recognize data subjects as having any proprietary interest in the data of which they are subjects. *See, e.g.,* Daniel J. Solove, *Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1880 (2013).

[23] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1509 (2019).

[24] Catherine Cheney, *How Facebook Statuses Informed the Zika Response in Brazil*, DEVEX (Dec. 13, 2016), https://www.devex.com/news/how-facebook-statuses-informed-thezika-response-in-brazil-89290.

[25] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1513 (2019).

[26] *Id.*

[27] These incentives are discussed in section II.c.ii below.

[28] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1513 (2019).

The Benefits of Data Sharing

(i)   Data as a Valuable Commodity

Effective public policy rests on good information about problems and the efficacy of the solution. While in the 19th and 20th centuries, governments obtained this information through statistics offices, a modality often touted currently is the analysis of big data, through which far more nuanced understandings of social problems can be gained.[29] Several qualities of big data have accelerated the growing demand of its use for the public good. First, there is an increasing demand for evidence-based social action[30] and the ability to capture and analyse real-time measurements that big data affords is especially valuable for social action that must quickly respond to unforeseen events.[31] Second, big data can reveal less conspicuous social action and public sentiment that may not be visible through traditional metrics, such as national statistics.[32] This is particularly the case for social media data, where individuals "team up" with other users to respond to social events, and express opinions on issues that they may not be so willing to share in an official forum. Finally, mobile data (while not the focus of this paper) is of particular value for humanitarian organizations because of its ability to reveal population movements in real time.[33]

(ii)   Private Sector Incentives

As alluded to above, data philanthropy arrangements offer a variety of incentives for private sector providers of data. While this paper presents a discussion of the risks inherent in data philanthropy to the organizations receiving the data, it is useful for such organizations to be aware of these private sector incentives so that they can better scrutinize proposals for data sharing. Klein and Verhulst suggest six categories of sharing incentives for private sector entities: reciprocity; research, recruitment and insights; reputation and public relations; increasing revenue; regulatory compliance; and responsibility and corporate philanthropy.[34] Of these, reciprocity, reputational advantage and corporate philanthropy are the most relevant to this discussion. As regards reciprocity, Klein and Verhulst outline one type that takes the form of compensation. Companies that collect personal information, particularly those in the social media industry, may act under the impression that they are giving back to society in order to counterbalance what they have taken from individuals, by way of their personal data.[35] As Lev-Aretz elaborates, this kind of data philanthropy serves as a form of non-mandatory tax, akin to the carbon tax, in that a company that pollutes society with surveillance "pays" by donating its data back for socially beneficial causes.[36] In relation to reputational advantage, it seems clear

---

[29] Rohan Samarajiva & Srinivas Lokanathan, *Using Behavioural Big Data for Public Purposes: Exploring Frontier Issues of an Emerging Policy Arena* 1, 3 (Feb. 1, 2016), https://lirneasia.net/wp-content/uploads/2013/09/NVF-LIRNEasia-report-v8-160201.pdf (last visited Nov. 19, 2019).

[30] Claudia J. Coulton et al., *Harnessing Big Data for Social Good: A Grand Challenge for Social Work* 4 (Am. Acad. of Soc. Work & Soc. Welfare, Working Paper No. 11, 2015), http://aaswsw.org/wp-content/uploads/2015/12/ WP11-with-cover.pdf.

[31] Rajesh Chandy, Magda Hassan & Prokriti Mukherji, *Big Data for Good: Insights from Emerging Markets*, 34 J. PROD. & INNOVATION MGMT. 703 (2017); Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1507 (2019).

[32] Peter Baeck (ed.), *Data For Good: How Big and Open Data Can Be Used for the Common Good*, NESTA 1, 25 (Feb. 2015), https://www.nesta.org.uk/sites/default/files/dataforgood.pdf.

[33] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1508 (2019).

[34] Thilo Klein & Stefaan Verhulst, *Access to New Data Sources for Statistics: Business Models and Incentives for the Corporate Sector* 17 (Paris21 Partnership in Stat. for Dev. in the 21st Century, Discussion Paper No. 10, 2017), http://www.thegovlab.org/static/files/publications/paris-21.pdf.

[35] *Id*, 12.

[36] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1511 (2019).

that sharing data for socially beneficial purposes can improve a business' reputation. In the wake of controversies concerning misuse of data, such as Facebook's Cambridge Analytical scandal, it seems that, now more than ever, companies wish to demonstrate that their collection and use of data practices align with the public interest.[37]

Modes of Data Sharing

In addition to mapping out different corporate incentive structures, Klein and Verhulst have also identified five different categories of data sharing: in-house production of statistics, transferring copies of data sets, remote access model, trusted third parties, and the algorithmic sharing model.[38] (1) *In-house production of statistics:* In the most common form of data collaboration, data holders process the data in-house and share the resulting statistics with problem holders, thus retaining control over the raw data. (2) *Transferring copies of data sets:* Data holders transfer copies of data sets that have been de-identified and aggregated, on which users (most likely skill holders) can run their algorithms. This method is most effective for research purposes because the aggregated data sets contain detailed information and a merger of different data sets and sources. However, producers relinquish control over the data and increase risks of security breaches and privacy harms.[39] (3) *Remote access model:* In this model, end users securely access data remotely while data holders maintain control over the extracted information. Data is not duplicated and does not leave the premises of the data holder, because users can only export the final aggregated metrics.[40] (4) *Trusted third parties:* Data holders and users rely on a trusted third party who possess reliable technical infrastructure, including large data storage capacity and secure connections, to facilitate secure access to the data.[41] Users do not have access to the raw data and instead request reports from the facilitator.[42] (5) *Algorithmic sharing model:* Finally, this model allows for data to remain in the data holders possession, and the data holder runs an algorithm of the data user's choice, with the data holder producing the final report.

---

[37] *Id*, 1511-1512.

[38] Thilo Klein & Stefaan Verhulst, *Access to New Data Sources for Statistics: Business Models and Incentives for the Corporate Sector* (Paris21 Partnership in Stat. for Dev. in the 21st Century, Discussion Paper No. 10, 2017), http://www.thegovlab.org/static/files/publications/paris-21.pdf.

[39] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1504 (2019).

[40] *Id*, 1505.

[41] Thilo Klein & Stefaan Verhulst, *Access to New Data Sources for Statistics: Business Models and Incentives for the Corporate Sector* 22-23 (Paris21 Partnership in Stat. for Dev. in the 21st Century, Discussion Paper No. 10, 2017), http://www.thegovlab.org/static/files/publications/paris-21.pdf.

[42] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1506 (2019).

PART TWO: CHALLENGES IN THE CONTEXT OF DEVELOPMENT POLICY

Value of the Data

Before turning to the privacy and other risks posed by the use of philanthropic data in policy development, it is first necessary to consider the actual value of the data itself. The use of big data can perpetuate inequalities and create new ones due to data bias and error. Unofficial data sources can be decentralized, unstandardized, unstructured and unrepresentive. The use of social media data, in particular, can result in marginalization, as we often falsely assume that the data is fairly representative of the general population and overlook the "data invisibles".[43] "Data invisibles" are those people who, generally due to their socio-economic status, are not counted or tracked within the digital economy.[44] They typically disproportionately include migrants, women, children, rural and slum dwellers who, often marginalized within their own societies, are not big producers or subjects of data.[45] Marginalization is particularly important in the policy arena, where action could be triggered by insights that have been developed on the basis of available data.[46] The question of representativity is particularly pertinent in the case of data generated in developing countries, where datafication is already incomplete. In fact, Samarajiva and others argue that in developing countries, it is better to rely on the simpler data generated by standard mobile networks (and not smartphones) to avoid marginalization of the poor.[47]

> *A risk inherent to using big data to monitor human rights is that it may privilege certain rights or outcomes that are better suited to quantitative measurement and fail to incorporate the relevant contextual information.*

Further, an additional risk inherent in using big data to monitor human rights (and, as this paper argues, development policy) is that it may privilege certain rights or outcomes that are better suited to quantitative measurement and fail to incorporate the relevant contextual information to shed light on the data.[48] In this way, objectivity through numbers and statistics is privileged over local contextual knowledge, and matters such as who is deciding what is counted as data, how it is interpreted, what measurements and indicators are used, and the purposes to which such data is put, are overlooked.[49] A data driven approach to human rights and development will not be useful for those areas where little data is generated or where there is insufficient

---

[43] Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 662, 669 (2016) ("Twitter does not represent 'all people', and it is an error to assume 'people' and 'Twitter users' are synonymous: they are a very particular sub-set. Neither is the population using Twitter representative of the global population. Nor can we assume that accounts and users are equivalent.").

[44] R. Shuman and F. Mita Paramita, *Why your view of the world is riddled with holes*, WORLD ECONOMIC FORUM (2016), www.weforum.org/agenda/2016/01/data-invisibles-ignore-at-our-peril/.

[45] Alberto Alemanno, *Big Data for Good: Unlocking Privately-Held Data to the Benefit of the Many,* 9 EUR. J. OF RISK & REG. 183, 184 (2018).

[46] Rohan Samarajiva & Srinivas Lokanathan, *Using Behavioural Big Data for Public Purposes: Exploring Frontier Issues of an Emerging Policy Arena* 1, 3 (Feb. 1, 2016), https://lirneasia.net/wp-content/uploads/2013/09/NVF-LIRNEasia-report-v8-160201.pdf (last visited Nov. 19, 2019).

[47] Rohan Samarajiva et al., *Big data to improve urban planning*, ECONOMIC AND POLITICAL WEEKLY 42-48 (May 30, 2015), http://www.epw.in/review-urban-affairs/big-data-improve-urban-planning.html.

[48] Galit A. Sarfaty, Can Big Data Revolutionize International Human Rights Law, 39 U. PA. J. INT'L L. 73, 90 (2017).

[49] SALLY ENGLE MERRY, THE SEDUCTIONS OF QUANTIFICATION: MEASURING HUMAN RIGHTS, GENDER VIOLENCE AND SEX TRAFFICKING, 4-6 (University of Chicago Press, 2016).

opportunity to identify trends over time.[50] The use of big data in the human rights and development fields is part of an increasing focus on quantification in global governance,[51] yet it risks over-reliance on data-driven approaches that exacerbate power imbalances and fail to protect vulnerable populations.

Privacy Concerns

(i)   Legal Risks

Data collected by private actors often contains or makes it easy to infer personal and sensitive details about individuals' lives. This is particularly the case for social media data. It is only natural, therefore, that discussions around the use of personal data for philanthropic purposes entail a discussion of the legal privacy risks. However, in reality, privacy compliance has not presented a significant barrier to corporate sharing for social good because most companies have shielded themselves from legal liability in two ways: (1) terms of service agreements and (2) technical and institutional privacy safeguards.[52] First, by broadly defining data collection purposes and granting authorization for academic research reuse, corporations collecting personal data have been able to contractually safeguard themselves from legal liability when sharing the data for socially beneficial purposes.[53] Thus, it appears that corporations who are truly interested in donating data have plenty of legal options to do so, and the extent to which legal privacy concerns hinder these efforts is overstated.[54] Second, to the extent that personal data is legally shared with third parties, privacy regulations impose certain safeguards in order to protect that data. Currently, global data protection law is broadly based on the Fair Information Practice Principles (FIPPs), which largely focus on individual privacy.[55] The gateway for governance by the FIPPs is the notion of Personally Identifiable Information (PII). In short, when an organization deals with PII, then the FIPPs apply, and if an organization does not deal with PII, then the FIPPs do not apply.[56] Thus, because PII is the gateway for most data protection laws, the easiest way to enable unregulated data sharing is to de-identify the data at issue. Accordingly, and as posited by Lee-Aretz, calls to provide a framework for privacy protections in data philanthropy tend to misidentify the risk and the required response, overemphasizing the need for legal regulation of data privacy.[57] While Lee-Aretz argues legal

---

[50] Galit A. Sarfaty, Can Big Data Revolutionize International Human Rights Law, 39 U. PA. J. INT'L L. 73, 90 (2017).

[51] *See, e.g.,* Kevin E. Davis et al., *Introduction: Global Governance by Indicators*, *in* GOVERNANCE BY INDICATORS 3, 12 (Kevin E. Davis et al., eds., 2012); Kevin E. Davis et al., *Introduction: The Local-Global Life of Indicators: Law, Power, and Resistance*, *in* THE QUIET POWER OF INDICATORS 1, 1 (Sally Engle Merry et al. eds., 2015).

[52] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1497 (2019).

[53] *Id,* 1495. *See, e.g.,* Facebook's Data Policy which states: "We transfer information to vendors, service providers, and other partners who globally support our business, such as providing technical infrastructure services, analyzing how our Services are used, measuring the effectiveness of ads and services, providing customer service, facilitating payments, or *conducting academic research and surveys*. These partners must adhere to strict confidentiality obligations in a way that is consistent with this Data Policy and the agreements we enter into with them." (Emphasis added). *Data Policy,* Facebook, https://www.facebook.com/policy.php (last visited November 3, 2019).

[54] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1495 (2019).

[55] *See, e.g.,* U.S. Dep't of Health, Educ. & Welfare, REP. NO. (OS)73-94 (1973) (recommending development of a Fair Information Practice code and highlighting its importance); EUR. PARL. DOC. (REG 679) (2016); EUR. PARL. DOC. (Repealing Directive 95/46/EC) (2016) ("GDPR") (recognizing a person's right to privacy in their data as a fundamental right); OECD, GUIDELINES GOVERNING THE PROTECTION OF PRIVACY AND TRANSBORDER FLOWS OF PERSONAL DATA (1980) (outlining guidelines for effective protection of data and privacy).

[56] Lisa M. Austin & David Lie, *Safe Sharing Sites,* 94 N.Y.U. LAW. REV. 581, 590 (2019).

[57] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1524 (2019).

recognition of the data philanthropy practice itself is indeed required,[58] this paper also calls for the development of ethical guidelines that recognize the normative risks and privacy harms that do not yet give rise to legal liability, yet are of little beneficial impact to the public good, which data philanthropy is meant to serve.

(ii)   Ethical Risks

a.   The Possibility of Re-Identification

Relying solely on de-identification to escape legal regulation for personal data transfers is a flawed strategy for two reasons. *First*, the boundary between PII and non-PII is unstable. Anonymization has historically been considered the ultimate cure for data privacy concerns, turning PII data into non-PII data. However, in the past two decades, a substantial body of research has shown that "anonymous" data can be re-identified in a number of ways.[59] In particular, researchers have been able to re-identify publicly released de-identified Netflix data,[60] AOL data,[61] and Australian Health Data.[62] Thus, researchers have repeatedly found that "anonymized" data is far from identification-proof. Once anonymized data is shared and aggregated, users can often be re-identified, because patterns of behavior recorded in the data are unique to one individual or because the anonymized data reveals an individual's identity once combined with other datasets.[63] While differential privacy, which measures the amount of privacy lost by individuals in a dataset,[64] has been offered as a mathematically driven solution to re-identification risks, some argue that it is still not identification proof.[65]

Second, strategies currently adopted to mitigate re-identification risks reduce the accuracy of the data.[66] If the goal of adopting a data-driven approach to development policy is to spur innovation and contemporary computational methods require large and accurate datasets, then there is an inherent tension between managing the risks of re-identification and ensuring data accuracy. Statistics Canada has tried to respond to this tension by developing its Research Data Centres (RDC) Program. While some research can be conducted using their de-identified Public Use Microdata Files, other forms of research that require access

---

[58] She argues that this can be achieved through an exception to the FIPPs that would reconcile data philanthropy with the purpose specification and use limitation principles. Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1524 (2019).

[59] *See, e.g.* Latanya Sweeney, *Simple Demographics Often Identify People Uniquely* (Carnegie Mellon Univ., Data Privacy Working Paper No. 3, 2000); Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization,* 57 UCLA L. REV. 1701 (2010); Lisa M. Austin, *Reviewing PIPEDA: Control, Privacy and the Limits of Fair Information Practices*, 44 CAN. BUS. L.J. 21, 35–36 (2006); Ira S. Rubinstein & Woodrow Hartzog, *Anonymization and Risk,* 91 WASH. L. REV. 703, 711–14 (2016); Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1836–48 (2011).

[60] Arvind Narayanan & Vitaly Shmatikov, *Robust De-Anonymization of Large Sparse Datasets,* 111 INST. ELECTRICAL & ELECTRONICS ENGINEERS SYMP. ON SECURITY & PRIVACY (2008).

[61] *See* Michael Barbaro & Tom Zeller Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, N.Y. TIMES (Aug. 9, 2006), https://www.nytimes.com/2006/08/09/technology/ 09aol.html.

[62] Chris Culnane et al., *Health Data In An Open World: A Report On Reidentifying Patients In The Mbs/Pbs Dataset And The Implications For Future Releases Of Australian Government Data* (2017), https://www.researchgate.net/ publication/321873477_Health_Data_in_an_Open_World.

[63] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1516 (2019).

[64] *See* Cynthia Dwork, *Differential Privacy*, 33 INT'L COLLOQUIUM ON AUTOMATA, LANGUAGES & PROGRAMMING, PART II, 1, 8–11 (2006) (defining differential privacy).

[65] Yafit Lev-Aretz, *Data Philanthropy*, 70 HASTINGS L.J. 1491, 1516 (2019).

[66] Lisa M. Austin & David Lie, *Safe Sharing Sites,* 94 N.Y.U. LAW. REV. 581, 583 (2019).

to more accurate datasets are conducted through the RDC Program, which allows users to access PII data within a secured facility.[67]

b. Collective Privacy

Data "is no longer gathered about one specific individual or a small group of people, but rather about large and undefined groups".[68] Increasingly, data analytics are focusing on pattern and group profiles which can be used to target people as members of a group.[69] For example, predictive policing algorithms are used to predict future neighborhood crime rates and allocate police resources, but have the effect of grouping people according to their racial profile.[70] Yet, ideas about group or communal privacy are typically missing from data protection regimes,[71] prompting some to argue that modern data protection policies should be amended to address these concerns.[72] In fact, some data protection policies developed by international organizations already recognize group privacy. For example, Médicins Sans Frontières and the World Health Organization policies make reference to the stigmatizing of communities themselves as a result of data collection.[73]

Group privacy concerns reflects the fact that data about people, even if de-identified, can raise numerous social, political and economic questions that need to be addressed by data governance frameworks.[74] For example, methods of handling data collected on Indigenous groups must be devised against the backdrop of colonialism, which raises various questions about where heritage materials are stored, how they are collected, by whom, and which community protocols should regulate access.[75] These issues are being discussed in the growing movement for "Indigenous data sovereignty", which considers the need for Indigenous groups to exercise a similar type of permanent sovereignty over their data as they would over their land and other natural resources.[76] These concepts are also present in the many data protection policies developed by First Nations Organizations which refer to "community privacy in the aggregate sense", a concept that appears to be quite foreign to many non-First Nations.[77]

---

[67] *See* THE RESEARCH DATA CENTRES PROGRAM, https://www.statcan.gc.ca/eng/rdc/ index (last visited Oct. 15, 2019).

[68] Linnet Taylor et al., *Introduction: A New Perspective on Privacy*, *in* GROUP PRIVACY: NEW CHALLENGES OF DATA TECHNOLOGIES 5 (Linnet Taylor et al. eds., 2017).

[69] *Id,* 5.

[70] *See, e.g.,* Jack Smith IV, *Crime-Prediction Tool May Be Reinforcing Discriminatory Policing*, BUSINESS INSIDER UK (Oct. 10, 2016); Alessandro Mantelero, *Personal Data for Decisional Purposes in the Age of Analytics: From an Individual to a Collective Dimension of Data Protection,* 32(2) COMPUTER L. & SEC. REV. 238, 240 (2016).

[71] Luciano Floridi, *Open Data, Data Protection, and Group Privacy*, 27(1) PHIL. & TECH. 1, 2 (2014) ("Such an "atomistic" ontology—take care of each member separately and the group will automatically be fine too—is at the roots of current European legislation.").

[72] Linett Taylor, Luciano Floridi & Bart van der Sloot (eds.), *Group Privacy: New Challenges of Data Technologies,* 126 PHILOSOPHICAL STUDIES SERIES 4-6 (2017), https://www.stiftung-nv.de/sites/default/files/group-privacy-2017-authors-draft-manuscript.pdf.

[73]*See* MÉDICINS SANS FRONTIÈRES, MSF DATA SHARING POLICY (Dec. 2013), https://fieldresearch.msf.org/handle/10144/306488; *see also* Unni Karunakara, *Data Sharing in a Humanitarian Organization: The Experience of Médicins Sans Frontières,* 10(12) PLOS MED. 1, 4 (2012); WHO, POLICY ON THE USE AND SHARING OF DATA COLLECTED IN MEMBER STATES OUTSIDE THE CONTEXT OF PUBLIC HEALTH EMERGENCIES: FREQUENTLY ASKED QUESTIONS at 4 ("In addition to the obligation to anonymize data . . . the policy includes measures to avoid stigmatization or exclusion of people or communities as a result of data collection.")

[74] Lisa M. Austin & David Lie, *Safe Sharing Sites,* 94 N.Y.U. LAW. REV. 581, 595 (2019).

[75] *Id*, 596.

[76] *See* TAHU KUKUTAI & JOHN TAYLOR, INDIGENOUS DATA SOVEREIGNTY: TOWARD AN AGENDA (2016) (discussing the need for data on indigenous people as well as the need for indigenous people to have control and sovereignty over that data); THE FIRST NATIONS PRINCIPLES OF OCAP®, https:// fnigc.ca/ocapr.html (last visited Apr. 27, 2019) ("OCAP® ensures that First Nations own their information and respects the fact that they are stewards of their information, much in the same way that they are stewards over their own lands.").

[77] *See, e.g.,* FIRST NATIONS INFORMATION GOVERNANCE CENTRE (FNIGC), OWNERSHIP, CONTROL, ACCESS AND POSSESSION (OCAP™): THE PATH TO FIRST NATIONS INFORMATION GOVERNANCE 34 (2014).

c. Informed Consent

In a "notice and consent" approach to privacy, there are often circumstances where individuals have unwittingly given their consent to the collection and use of web-generated data by ticking a box without reading the fine print.[78] In fact, relevant scholarship reveals that users of online services often do not read disclosures, or at best skim-read them.[79] There are a variety of reasons to explain this practice. Users may have a lack of interest in understanding what they are consenting to, face a difficulty understanding the legal language, or simply find reading these contracts to be too time consuming to properly digest.[80] Further, consumers are aware of their lack of bargaining power in negotiating these terms of service, the inconvenience of not accessing online services in an increasingly digitized society, and assume that because a great number of consumers are bound by the same terms, they must be reasonable.[81] However, as many advocates note, this kind of technical consent rarely resembles a true, clear and informed consent.[82] Thus, in a world where standard form contracts and online terms of service allow corporations to collect data for very broadly defined purposes,[83] "serious questions arise… about the extent to which consumers are truly aware of what data they are sharing, how and with whom, and to what use they will be put."[84] Further, once consent is given, it may be difficult or impossible for individuals to revoke, as there may be no clear channel through which to do so, or it may no longer be possible to erase the data.[85] Thus, using data-driven approaches to monitor human rights and development outcomes risks paradoxically causing further human rights violations.

Sustainability and Capacity Building

Data philanthropy partnerships face various operational challenges that prevent the transformative potential of data collaboration from being achieved.[86] Oftentimes, data collaboratives are short-term projects, making it difficult to establish more sustainable legal bases for effective collaborations.[87] Further, there is an alarming data asymmetry between the private and public sectors.[88] Private data, particularly social media data, remains the prerogative of a few very big corporations, and thus there is little incentive for corporations to release data into the public sphere and lose their competitive advantage.[89] As a result, the demand side lacks bargaining power, as parties do not all have the same capabilities, resources or experience when it comes to

---

[78] Galit A. Sarfaty, Can Big Data Revolutionize International Human Rights Law, 39 U. Pa. J. Int'l L. 73, 89 (2017).

[79] See, e.g., Ian Ayres & Alan Schwartz, The No-Reading Problem in Consumer Contract Law, 66 Stan. L. Rev 545, 547 (2014).

[80] Geslevich Packin & Lev-Aretz, Big Data and Social Netbanks: Are You Ready to Replace Your Bank?, 53 Hous. L. Rev. 1211, 1279 (2016).

[81] Yafit Lev-Aretz, Data Philanthropy, 70 Hastings L.J. 1491, 1517 (2019).

[82] See, e.g., Daniel J. Solove, Privacy Self-Management and the Consent Dilemma, 126 Harv. L. Rev. 1880 (2013).

[83] Julie E. Cohen, Examined Lives: Informational Privacy and the Subject As Object, 52 Stan. L. Rev. 1373, 1432-35 (2000); Jeff Sovern, Opting In, Opting Out, or No Options at All: The Fight for Control of Personal Information, 74 Wash. L. Rev. 1033, 1072-74 (1999); Richard Warner, Undermined Norms: The Corrosive Effect of Information Processing Technology on Informational Privacy, 55 St. Louis L. J. 1047, 1084-86 (2011); Tal Zarsky, Mine Your Own Business, 5 Yale. J. L. & Tech. 1, 33 (2002-2003); Katherine J. Strandburg, Free Fall: The Online Market's Consumer Preference Disconnect, U. Chi. Legal F. (2013).

[84] U.N. High Commissioner for Human Rights, The Right to Privacy in the Digital Age, U.N. Doc. A/HRC/27/37, at 18 (2014).

[85] Galit A. Sarfaty, Can Big Data Revolutionize International Human Rights Law, 39 U. Pa. J. Int'l L. 73, 89 (2017).

[86] Andrew Young, Andrew J. Zahuranec, Stephen Burley Tubman, William Hoffman & Stefaan Verhulst, Addressing the Challenges of Drafting Contracts for Data Collaboration, Data & Policy (Oct. 23, 2019).

[87] Stefaan Verhulst, How to Use Data for Good — 5 Priorities and a Roadmap, Medium (2019), https://medium.com/data-stewards-network/how-to-use-data-for-good-5-priorities-and-a-roadmap-df96c3477abc.

[88] Alberto Alemanno, Big Data for Good: Unlocking Privately-Held Data to the Benefit of the Many, 9 Eur. J. of Risk & Reg. 183, 185 (2018).

[89] Galit A. Sarfaty, Can Big Data Revolutionize International Human Rights Law, 39 U. Pa. J. Int'l L. 73, 91 (2017).

negotiating data sharing agreements.[90] Problem holders need to be wary of entering into asymmetrical contracts that effectively transform corporations into the primary gatekeepers of the data that protects human rights.[91] In addition, they must ensure that there are sufficient contractual mechanisms in place to address any repudiation on the supply side that may significantly affect the work of the organizations receiving the data.

One approach to increasing the capacity of problem holders to contract for data collaboration is to engage in multi-stakeholder discussions in order to better understand the needs, opportunities, challenges and risks related to data collaboration from both the supply and demand side. Contracts for Data Collaboration (C4DC) is one such initiative that seeks to identify and address various barriers to data collaboration.[92] Participants at a C4DC workshop in September 2019 agreed on the importance of creating an evidence base of current practices in the drafting of legal contracts for data collaboration (including the creation of a shared repository of template agreements), a glossary of key data collaboration terms, and a standard engagement and communication strategy in order to steer data collaboration contracting toward greater effectiveness and legitimacy.[93]

---

[90] GovLab, *Data: The Lever to Promote Innovation in the EU*, Medium. (2019), https://medium.com/data-stewards-network/data-the-lever-to-promote-innovation-in-the-eu-a1d13404698d.

[91] Galit A. Sarfaty, *Can Big Data Revolutionize International Human Rights Law*, 39 U. Pa. J. Int'l L. 73, 88 (2017).

[92] Contracts for Data Collaboration, https://contractsfordatacollaboration.org/ (last visited Oct. 9, 2019).

[93] Andrew Young, Andrew J. Zahuranec, Stephen Burley Tubman, William Hoffman & Stefaan Verhulst, *Addressing the Challenges of Drafting Contracts for Data Collaboration,* Data & Policy (Oct. 23, 2019).

PART THREE: MOVING FORWARD

A Trusted Data Sharing Platform

Trust is needed between the actors in a data collaborative and the general public, particularly in light of the many recent stories about the misuse of data. Stakeholders can build trust between parties by developing international standards that move beyond the bare minimum legal obligations, and take into account some of the normative concerns discussed in this paper. One

International standards, beyond the bare legal minimum obligations, are needed to take into account normative and ethical concerns about data philanthropy and build trust between parties in a data collaborative.

mechanism within the United Nations (UN) system in which this could occur is the UN Global Platform for Data, Services and Applications. Relevantly, this platform delivers a so-called "trusted" environment for collaborative data-analysis in order to help countries measure the Sustainable Development Goals (SDGs). The Platform is based around four pillars: trusted partners, trusted data, trusted methods (algorithms) and trusted learning. According to the Platform, trust is to be achieved through collaboration, peer review and approval of all work undertaken.[94] However, the Platform's concept of "trust" does not appear to be understood in ethical terms. Further, the datasets available on the Platform do not presently include PII that would

give rise to the ethical concerns discussed in this paper, including marginalization, re-identification, collective privacy and informed consent.[95] Therefore, while the Global Platform could provide a useful pre-existing digital infrastructure through which to share data containing PII with the UN system, it would not address the normative risks associated with data philanthropy. To that end, the UN system needs to establish a set of agreed ethical principles that it would apply to any PII data received from third parties. Only by incorporating those ethical guidelines into the Platform's notion of "trust", could the infrastructure serve as the appropriate mechanism through which to ethically incorporate PII data into the SDG monitoring process.

Safe Sharing Sites

One tension discussed in this paper is the extent to which data accuracy should be balanced against the need to protect against privacy risks associated with PII data. Professor Lisa M. Austin has outlined a method through which stakeholders can resolve this tension in ways that suit their individual needs: the Safe Sharing Site (SSS).[96] In so doing, Austin references a common scenario: Organization A has data that Organization B wishes to use. The SSS architecture would, in its simplest terms, allow Organization A to share data with Organization B without affording access to its raw form data, while providing Organization B

---

[94] *See* UN Global Working Group on Big Data for Official Statistics, *United Nations Global Platform: Data for the world* (2019), http://publications.officialstatistics.org/assets/pdf/UNGlobalPlatform_Brochure.pdf.

[95] Rather, they comprise: the *ABS-B Data Feed*, the world's largest source of unfiltered flight data; *Sentinel-1*, a pair of European synthetic aperture radar (SAR) satellites that produce sea and land monitoring data; *Sentinel-2*, a land-monitoring constellation of two satellites that provide high resolution optical imagery and continuity for the current SPOT (Satellite Pour l'Observation de la Terre) and Landsat missions; *Nielsen Scanner Data*, which contains weekly detailed barcode level data for certain product groups; and African Soil Information Service Soil Chemistry data, NOAA Global Historical Climatology Network Daily Data, and OpenStreetMap data. *See* UN Global Working Group on Big Data for Official Statistics, *United Nations Global Platform: Data for the world* 17-22 (2019), http://publications.officialstatistics.org/assets/pdf/UNGlobalPlatform_Brochure.pdf.

[96] Lisa M. Austin & David Lie, *Safe Sharing Sites,* 94 N.Y.U. LAW. REV. 581, 581 (2019).

with more accurate data results. Essentially, rather than providing Organization B with heavily de-identified data that safeguards against any type of computation that Organization B might perform, Organization A would instead transfer its raw form data into a secure SSS. Then, the data would be de-identified but only for the purpose of Organization B constructing a specific computation. However, once Organization B's computation is constructed, it would be performed on the raw dataset within the confines of the secure SSS. Without an SSS, Organization B would have to perform the computation itself on the de-identified dataset and achieve lower precision because of the all-encompassing privacy protecting technology. Yet, because of its modularity, an SSS is not a standalone solution to issues of privacy compliance.[97] However, its architecture could be employed in conjunction with, for example, the UN Global Platform, in order to ensure that the data upon which development policies are based is as accurate as it can be, while safeguarding the privacy of the data subjects.

---

[97] *Id*, 617.

CONCLUSION

The use of commercial data for the public good has enormous life-improving potential,[98] and is becoming an increasingly more common practice. Real-time data can assist immeasurably in a crisis situation as it can pinpoint the exact location of individuals in need of assistance. However, it can also tell us whether policy interventions are working in practice. As a result, international organizations, civil society groups and public sector entities are increasingly turning to partnerships with the private sector in order to use private sector data for the public good. However, in the rush towards innovation, many of the risks inherent in data sharing arrangements are overlooked. Thus, this paper provides a normative overview of the risks that should be considered before entering into such arrangements. In order to provide a more fruitful contribution, this paper focuses only on arrangements involving the transfer of personal data for the purposes of further development policy, and discusses these risks from the perspective of the data receivers.

In so doing, this paper first unpacks the notion of data philanthropy, including different modes for sharing, the role that each actor plays in a sharing arrangement, and the self-defining notion of what is considered "good" in data philanthropy. This paper calls on data receivers to negotiate agreements with an understanding of the functionalities of each modality of sharing arrangements, while remaining conscious of the interests of the corporate sector that are served by data philanthropy, the need to enter into sustainable relationships, and the self-defining nature of the social causes that are supposedly served by these arrangements.

Subsequently, the paper calls on data receivers to consider the value itself of the data they receive when applied to their specific development needs. Data, particularly social media data, can be non-representative and overlook "data invisibles". While such flaws might be considered justified in an emergency situation, where the focus is on saving lives at any cost, these issues become less palatable when the data is used for longer term policy in respect of which data receivers are arguably more accountable.

This paper then conducts a detailed examination of the various ethical privacy risks to data sharing that are not presently addressed by legal regimes. While outlining the legal distinction between PII and non-PII data in order to frame the legal structures that allow for data sharing, this paper agrees with Lev-Aretz's position and posits that legal privacy compliance does not function as a real barrier to data collaboration. Instead, it focuses on the normative privacy risks such as the risk of re-identification, collective privacy and informed consent. It does this in order to highlight a paradox in data collaboration: how can an international development organization, such as the United Nations, accept data from private companies in circumstances where the private sector's data collection practices do not align with the organization's position on ethics and privacy?[99] To this

---

[98] Alberto Alemanno, *Big Data for Good: Unlocking Privately-Held Data to the Benefit of the Many,* 9 EUR. J. OF RISK & REG. 183, 184 (2018).
[99] *See, e.g.* U.N. High Commissioner for Human Rights, *The Right to Privacy in the Digital Age*, U.N. Doc. A/HRC/27/37, at 18 (2014).

end, the paper calls for the development of a set of ethical guidelines for data collaboration that international organizations can adopt and apply to their external partnerships.

Finally, this paper offers two potential solutions that would help address some, but not all, of the risks discussed in this paper. The first of these is the UN Global Platform for Data, Services and Applications, which could provide a useful existing mechanism by which data is shared between the UN system and external private sector companies. Second, this paper deals with the concept of the Safe Sharing Site, which would also serve as a possible infrastructure to help alleviate some of the re-identification concerns while limiting the impact on data accuracy. It is very much stressed that these are not all-inclusive solutions, and are merely being offered to assist with some of the discrete issues present in the practice of data philanthropy collaborations, and brainstorm ideas for moving forward with collaborations. It remains the case that data receivers must develop a set of ethical guidelines that can be applied to data philanthropy proposals, such that the data used for development purposes is gathered in a manner consistent with the values for which a development organization advocates.