

# From FAIR data to fair data use: Methodological data fairness in health-related social media research

Big Data & Society  
January–June: 1–14  
© The Author(s) 2021  
DOI: 10.1177/20539517211010310  
journals.sagepub.com/home/bds  
SAGE

Sabina Leonelli<sup>1</sup> , Rebecca Lovell<sup>2</sup>, Benedict W Wheeler<sup>2</sup>,  
Lora Fleming<sup>2</sup>  and Hywel Williams<sup>3</sup> 

## Abstract

The paper problematises the reliability and ethics of using social media data, such as sourced from Twitter or Instagram, to carry out health-related research. As in many other domains, the opportunity to mine social media for information has been hailed as transformative for research on well-being and disease. Considerations around the fairness, responsibilities and accountabilities relating to using such data have often been set aside, on the understanding that as long as data were anonymised, no real ethical or scientific issue would arise. We first counter this perception by emphasising that the use of social media data in health research can yield problematic and unethical results. We then provide a conceptualisation of *methodological data fairness* that can complement data management principles such as FAIR by enhancing the actionability of social media data for future research. We highlight the forms that methodological data fairness can take at different stages of the research process and identify practical steps through which researchers can ensure that their practices and outcomes are scientifically sound as well as fair to society at large. We conclude that making research data fair as well as FAIR is inextricably linked to concerns around the adequacy of data practices. The failure to act on those concerns raises serious ethical, methodological and epistemic issues with the knowledge and evidence that are being produced.

## Keywords

Fairness, data mining, data ethics, research practice, social media, injustice, data use

## Introduction

There is growing agreement among scientists and policy makers that data used for research purposes, including those extracted from social media, should be made accessible whenever possible; and that their re-use should be encouraged, provided that it conforms with legal and ethical concerns around privacy and personal data protection, such as championed through the European General Data Protection Regulation (GDPR, 2019). To participants in the Open Data movement, this position is exemplified by the motto ‘as open as possible, as closed as necessary’.

Following years of debate on the values and practices underpinning data sharing and re-use, the FAIR principles have emerged as plausible guides towards the handling of data for research purposes (Mons et al., 2017; Wilkinson et al., 2016). Accordingly, data should be organised and managed to be easily *Findable*; *Accessible* to as many as possible, in ways that are user-friendly and machine-readable; *Interoperable* to foster

links with other data; and *Reusable*, i.e. easy to re-purpose. The FAIR principles have become a reference point for data management and big data infrastructures around the globe, such as the European Open Science Cloud – the platform set up by the European Commission to coordinate the stewardship of research

<sup>1</sup>Department of Sociology, Philosophy and Anthropology, Exeter Centre for the Study of the Life Sciences (Egenis), University of Exeter & Alan Turing Institute, Exeter, UK

<sup>2</sup>European Centre for Environment and Human Health, University of Exeter Medical School Knowledge Spa, Royal Cornwall Hospital, Truro, Cornwall, UK

<sup>3</sup>Computer Science, University of Exeter & Alan Turing Institute, Exeter, UK

### Corresponding author:

Sabina Leonelli, Department of Sociology, Philosophy and Anthropology, Exeter Centre for the Study of the Life Sciences (Egenis), University of Exeter & Alan Turing Institute, Byrne House, St Germans Road, Exeter EX4 4PJ, UK.

Email: s.leonelli@exeter.ac.uk



data (High Level Expert Group European Open Science Cloud, 2016). They have been recently complemented by another two sets of principles aimed at considerations of data ethics and responsible use, which were not explicitly included in the original FAIR formulation: (1) the CARE principles for indigenous data governance, which include Collective benefit, Authority to control, Responsibility and Ethics ([www.gida-global.org/care#](http://www.gida-global.org/care#)) and aim to safeguard indigenous knowledge systems from predatory mining and appropriation; and (2) the TRUST principles, which include Transparency, Responsibility, User focus, Sustainability and Technology (Lin et al., 2020) and aim to enhance accountability in the development and maintenance of data infrastructures.

While applauding these developments, this paper identifies a research area in which the application of FAIR, CARE and TRUST principles is not enough to guarantee that data collection, processing and use are fair to those affected by these processes. This is the use of social media data (SMD) for health-related research.<sup>1</sup> As in other domains, the opportunity to mine social media for information has been hailed as transformative for research on human health, given its power to document human behaviour in real time. Considerations around the fairness and accountabilities relating to using SMD have often been set aside, on the understanding that as long as data were anonymised, no substantive issue would arise. We counter this perception by showing that the use of SMD in health research can yield results that are both scientifically dubious and ethically problematic. Our argument is that the use of SMD, no matter how well-intentioned and/or compliant with FAIR principles, is not necessarily fair to those from whom the data are gathered and those whose lives may be affected by specific uses of the data.

This problem is only indirectly addressed by the CARE and TRUST frameworks, since it does not necessarily concern indigenous data sources (the main focus of CARE) and it is not limited to the set-up of data infrastructures (the main targets of TRUST). Rather, social media capture data from individuals around the world, including highly resourced countries – indeed, the fact that they may disproportionately capture information from specific demographic, social or cultural groups constitutes a key concern around their use for research. Moreover, concerns around the social and ethical implications of reusing such data for health research affect the whole research process, including the design and use of databases as well as models, algorithms and the choice of directions and goals.

To develop this argument, Using social media as research data section of this paper discusses how uncritical uses of SMD in health research risk fostering

existing social divides, promoting misrepresentations of society and breaking public trust in research institutions – with damaging effects on research outputs, applications and public perception. What is data fairness within research? section considers what fairness means in the context of SMD use, emphasising its significance towards guaranteeing the reliability of knowledge produced through data interpretation. We propose to focus on *methodological data fairness* as a form of fairness that complements data use principles such as FAIR by enhancing the actionability of SMD for future research. Section 3 exemplifies how this focus on methodological data fairness informs different stages of the research process, including study design, data collection, storage, processing and interpretation. We identify practical steps through which social media researchers can foster practices and outcomes that are methodologically sound and fair to society at large.

Our contribution to wider debates on data ethics and justice is thus to identify the ethical problems arising from scientists' uses of SMD, and to suggest that these problems may be addressed by incorporating a commitment to fairness in the methods routinely used in health research. We do not aim to provide a comprehensive review of what data justice and fairness may involve across contexts, nor do we aim to offer an empirical analysis of the inequity and discrimination affecting social media platforms, which has already been provided by many scholars including Srnicek (2017), O'Neill (2017), Zuboff (2019) and Noble (2018). Rather, we build on such scholarship to reflect on what researchers using SMD can concretely do to counter such bias. Our suggestions are thus complementary to calls for increased responsibility and accountability by social media providers (Dijck et al., 2018; Lupton and Michael, 2017; O'Neill, 2017; Van Kleek et al., 2017) and scholarship on the intersections between platform capitalism and social studies of infrastructures (Plantin et al., 2018).

Our analysis draws from relevant literature in data studies, health research and philosophical debates on fairness, as well as our experiences as a UK-based interdisciplinary group engaging in data-intensive, health-related research from the perspective of computer science (HW), public health (RL, LF, BW) and critical data studies (SL). Between 2016 and 2017, we had opportunity to work together within the project 'Social sensing of health and wellbeing impacts from pollen and air pollution', resulting in the present exploration of the risks and opportunities involved in SMD analysis. Accordingly, we will accompany our argument with illustrations of how our research was informed and/or constrained by methodological fairness.

## Using social media as research data

Advances in data availability and analysis have outpaced the process of ensuring that data-intensive research is equitable and appropriate with respect to the publics affected by its results. Research in social science and the humanities has shown how mining personal data can lead to discrimination and damage to vulnerable communities (boyd and Crawford, 2012; D'Ignazio and Klein 2020; Floridi, 2011; Floridi and Taddeo, 2016). In the medical sciences, there are related concerns around handling sensitive information (Auffray et al., 2016; Dijck and Poell, 2016; Tamar and Lucivero, 2019). The ease of use fostered by the FAIR principles can be misunderstood to encourage disregard for the ethical implications of research. Furthermore, big data analysis can suffer from design and methodological problems, such as inadequate theory development, lack of sampling methods and justification, limited control over and information about data sources and difficulties in assessing the extent to which a given dataset is representative of a phenomenon of interest (Metcalf and Crawford, 2016, Taylor, 2017, Veale and Binns, 2017). These issues can compromise the credibility, reliability, and applicability of findings.

Social media offer the researcher a rich and extensive source of data at a scale not previously achievable (Kavanaugh et al., 2012; Leavey, 2013; Picazo-Vela, et al., 2012). Several research sectors (including business and government) seek to exploit the potential of the vast amounts of information generated through social media. The processes associated with the accessing, transformation, analysis and interpretation of SMD need to be critically assessed. As Halford et al. (2017) argued, 'far from being "naturally occurring"...social media data are shaped through complex social and technical processes'. These processes have societal implications that pose specific regulatory challenges (Pentzold and Fischer, 2017). When SMD analysis is robust and reliable, the resulting knowledge can enhance understanding of countless topics, including some which have traditionally been difficult to address, and counterbalance other methods of data collection (such as social surveys) that can also over- or under-represent certain groups. Further, the use of SMD offers opportunities to reform relationships between governments and citizens, enhance democratic representation and inform public policy and decision-making (Leavey, 2013; Picazo-Vela et al., 2012) – for instance, by improving service delivery; developing early warning indicators of key societal issues (Hays and Daker-White, 2015; Leavey, 2013) and guiding responses to disaster situations (Kavanaugh et al., 2012).

Critical scholarship on the use of social media has initially focused on privacy and surveillance (Conway, 2014; Mikal et al., 2016), with attention to social fairness emerging more recently (Kennedy et al., 2017; Taylor, 2017). Opaque and unfair social media research practices may covertly embed and exacerbate structural social disparities in power, magnifying social divides, inequalities and injustices; and may promote misrepresentations of societies, communities and individuals (D'Ignazio and Klein, 2020; Floridi, 2011; Munoz et al., 2016). There is also potential for damaging effects to the reputation and public perception of science.

Public concerns regarding SMD extend beyond the protection of privacy to an interest in whether social media and data analytics companies are meeting wider expectations of what is fair (Kennedy et al., 2017). As demonstrated by the Cambridge Analytica scandal around Facebook data misuse, high-profile controversies regarding inappropriate and unethical use of SMD threaten the perceived trustworthiness of institutions such as universities and governmental bodies (O'Neill, 2004; Pentzold and Fischer, 2017). Any backlash against the use of such data thus risks to compromise the valuable insights that these data can provide (Halford et al., 2017).

The question of what is appropriate when using SMD is 'urgent as data and research agendas move well beyond those typical of the computational and natural sciences, to more directly address sensitive aspects of human behavior, interaction, and health' (Zook et al 2017). This urgency was underlined by Taylor and Pagliari's (2017) finding that only one of the UK research councils had ethical guidance for social media research (with two others pointing to guidelines by other bodies); and only 50 out of 156 studies mentioned ethics (including statements that no approval was required) – a figure likely to be lower if restricting the sample to computer science, where ethical considerations are not always considered and related oversight is often lacking. This is despite many recent efforts to identify procedural and ethical frameworks for the gathering, transformation, analysis and dissemination of SMD (Dove et al., 2016; Gelinis et al., 2017; Golder et al., 2017; Lee, 2017; Mittelstadt and Floridi, 2016; Townsend and Wallace, 2016; Weller and Kinder-Kurlanda, 2016). Solutions to embedding ethical principles in social media research are lagging behind due to the pace of development and change in the technologies available, the widening of disciplinary interests as the potential of the data becomes clear and the contextual variability of the notion of fairness in relation to social media research use. Whilst there is public support for the fair use of SMD for health research, this appears

dependent on the methods used and adequate protection of privacy (Mikal et al., 2016).

Perceived motivations of and trust in institutions have also emerged as salient factors in public perceptions of fair use (Williams et al., 2017). Practices which would be unacceptable for commercial goals may be considered differently in case of demonstrable public gain. Identifying the expectations of social media users regarding what is fair is complicated by the complexity of SMD and what they represent, and by low levels of public understanding regarding who owns and has access to the data, and how they can be used (boyd and Crawford, 2012; Kennedy et al., 2017; Nissenbaum, 2009). Many social media users do not know who is accountable for the secondary use of their data (boyd and Crawford, 2012).

There are questions as to whether the protections put in place through legal frameworks such as the GDPR are flexible and responsive enough to ensure fairness, especially in the face of rapidly developing methodologies and applications of SMD (Halford and Savage, 2017). The contextual dependency of fairness, difficulties with identifying what is generally acceptable, variability around what is appropriate between contexts and the often conflicting interests of different groups or of different goals, all present challenges to Institutional Review Board or Research Ethics Committee assessments of SMD use (Hunter et al., 2018; Zook et al., 2017). By exercising data fairness at each stage of the research process section, we argue that careful planning and management of research can help overcome these challenges. Before delving into practical considerations, however, we need to add conceptual clarity to the notion of data fairness by introducing the idea of '*methodological data fairness*'.

### What is data fairness within research?

In the Oxford English Dictionary 'fairness' is defined as 'the quality of treating people equally or in a way that is right or reasonable'. It is clear from this definition that identifying what constitutes fairness is complicated by the context dependency of what is right or reasonable (Bicchieri and Chavez, 2010; Wolff, 2010), which in turn relates to what one is trying to achieve (Ryan, 2006). It is thus no surprise that notions of trust (O'Neill, 2004), accountability (boyd and Crawford, 2012), transparency (Information Commissioners Office, 2017) and justice (Taylor, 2017), all come into play.

Many existing analyses of fairness focus on the extent to which goods or opportunities are available to different individuals or groups, and whether existing disparities in the allocation of resources can be

alleviated. This *distributive understanding of fairness* has been applied to discussions around access to public goods, and issues of the environment and sustainability, education and social justice (Cullity, 2008). It has also been considered in relation to the rights and responsibilities of accessing, gathering, transforming and using big data (Kennedy et al., 2017; Nissenbaum, 2009; O'Neill, 2017). This scholarship crucially complements the implementation of the FAIR principles, which despite their misleading acronym do not address the benefits and harms arising from data management and re-use. Compliance with the FAIR principles may well make a given dataset available for computational mining, but it does not question the effects of such accessibility on data subjects and communities impacted by data-driven decision-making systems.

This same literature on distributive fairness has acknowledged that it is hard to determine equal distributions of data or equal stewardship of data, since what constitutes equal – and who is taken into account among data subjects and analysts – depends on the situation of inquiry. For example, our interdisciplinary project aimed to analyse SMD to monitor the prevalence and location of hayfever symptoms across the UK. We used Twitter data seeking mention of hayfever symptoms to see if this would provide early warning of the effects of increased pollen levels. We successfully obtained ethics approval for this research, however none of the people who sent these tweets had been informed or consented to the use of their tweets for this purpose, including the people whose actual geographic location was identifiable at the time of sending the tweets. This was a relatively benign use of 'anonymised' data but was it a fair use? What would the equal distribution of data mean in this context? How does the lack of contextual information about the users other than their location (such as their age, gender and socio-economic status) affect research outputs? And if Twitter users were aware of such data reuse, would they change the ways in which they discuss hayfever on the platform, and what implications would this have for our research?

Questions of this kind emerge whenever SMD are used for knowledge production, and become particularly sensitive in relation to health research, with social media users expressing worries about potential discrimination when asked about the rights and wrongs of using their data for research (Kennedy et al., 2017). To address such issues, we need an understanding of fairness that places greater emphasis on data usage and its relation to individual and collective agency within society. Consideration of fairness needs to actively foster the use of data to produce knowledge that supports social agency and empowers vulnerable groups,

thus explicitly targeting concerns around visibility and representation. This is where the concept of *data justice* proves effective, focused as it is on fostering fairness in the ways in which people are ‘*made visible, represented and treated*’ through big data (Galbraith, 2017; Taylor, 2017), thus enhancing people’s capacity for action. Data justice aims to counter undocumented bias in datasets that over-represent some population to the expense of others, thus actively and explicitly identifying and eradicating forms of discrimination that limits people’s ability to participate in society (D’Ignazio and Klein, 2020; Veale and Binns, 2017). This is especially relevant in relation to the use of SMD as evidence for scientific claims.

In what follows, we aim to build on such literature and our own insights as data practitioners to answer a critical question for knowledge production activities grounded on the analysis of SMD: how can such research treat data fairly in ways that enable positive action? In other words, can specific practices and methods of SMD analysis foster distributive fairness and data justice? We propose to think broadly about the role of data within epistemic processes (i.e. processes relating to the conditions for knowledge creation) and to consider the forms of injustice tied to the use of data within research, thus linking *methodological* considerations to the social context in which methods are used.

In this respect, the work of philosopher Miranda Fricker, though not specifically devised to address data management strategies, provides a helpful starting point. Fricker (2009) argues that evaluations of what is fair and just within society involve examining the forms of prejudice built into our ways of understanding the world (which can include tools, methods, knowledge claims and of course, data). She considers the extent to which our ways of knowing enable us to treat people in ways that are right or reasonable, and proposes to distinguish between two forms of epistemic injustice: (1) *testimonial injustice*, defined as prejudice in the existing economy of credibility (e.g. there is a general prejudice against black witnesses being reliable, and therefore the police may not believe statements made by a black person); and (2) *hermeneutical injustice*, defined as prejudice in the economy of collective resources used to make sense of the world (e.g. in a culture in which there is no concept of sexual harassment and no legal recognition of this as a crime, it is not possible for victims to bring attention to their plight).

In the world of data practices, examples of testimonial injustice appear in the choice of social media platforms as preferred data source, due to those platforms being the easiest and least expensive to access. A clear example is Twitter, which combines a relaxed approach to data re-use with the opportunity to download and

re-use for free a sample of the data available on the platform on any given day. For publicly funded projects such as ours, the absence of upfront costs in accessing the data is a big carrot, which explains the predominance of Twitter as a data source for health-related research based on social media (Sinnenberg et al., 2017). This predominance makes it difficult to question the use of Twitter as a data source in the first place, despite its shortcomings; and to argue for alternative data sources requiring additional investment. This has created a prejudice in favor of Twitter data use within health research, with methods, software and training tools emerging that targeted specifically this data source and made it even easier to deploy, especially for newcomers to the field (Sinnenberg et al., 2017).

Hermeneutical injustice takes the form of privileging specific socio-demographic, cultural or political stances within data sources, as in the charged cases of migration or obesity figures. Consider studies that use Twitter data for health research without taking account of the types of people who do and do not use this platform, and thus of the bias built into their data source. Hermeneutical injustice happens whenever a published study mentions the fact that Twitter users are largely young, in professional occupations, and based in cities rather than rural areas, without however taking those biases into account in the rest of the analysis. Some studies may argue that such biases are not relevant in their case, e.g. when investigating availability and access to health care resources across a region. Attention to hermeneutical injustice however demands that assumptions made around whether or not a given dataset is representative, and of what, be examined empirically and in detail. What evidence is there that data collected on young urban dwellers represent the views and experiences of older people living in rural areas? And what are the implications of this set-up for the studies at hand?

The concepts of hermeneutical and testimonial injustice expose the significance of focusing on research methodology as a key locus for decisions around what constitutes data, for which purposes and audiences, and with which social and ethical implications. Perhaps surprisingly, researchers working in data science and related domains are not necessarily thinking of fairness as a factor underpinning their everyday technical choices and methods. Data fairness is often viewed as a legal and regulatory issue, whose resolution is in the hands of judiciary and policy bodies. This understanding of data fairness focuses primarily on consent and protection from harm, though legislation around equality of opportunity and unlawful discrimination is also relevant to how researchers use personal data (Clifford and Ausloos, 2018). Researchers are required to comply with this legal understanding of

data fairness, which provides broad guidance on what to avoid when handling SMD (such as the publication of easily identifiable, sensitive information without the consent of the individuals concerned). However, legal frameworks do not provide concrete guidance on all aspects of research practice, especially since what legal precepts mean for everyday technical decisions unavoidably depends on the context of application.

A crucial complement to the legal approach is a social understanding of data fairness as a way to counter discrimination and decrease social divides (Fiesler et al., 2020; Veale and Binns, 2017; Zook et al., 2017). This involves considering the intention and approach of data practices, the extent to which different publics and stakeholders are engaged, the power relations at play and the role played by trust in facilitating or disrupting data access and mining. Though familiar to researchers who use big data, these considerations are still viewed as largely extraneous to research itself (Golder et al., 2017). Indeed, making data fair is often delegated to governance mechanisms (such as ethics committees, review boards and public engagement groups) that encourage awareness of potentially discriminatory mechanisms among researchers and wider society. These mechanisms, however, may fall short of providing guidance to researchers when it comes to technical data practices ranging from data cleaning to storage and modelling.

We propose that what is missing is an understanding of data fairness as part and parcel of the research process. *Methodological data fairness* focuses on how the treatment of data during the course of a study affects the credibility and justice with which the outputs of the study portray and affect all sections of society. Here the main concern is with the quality and credibility of the research processes through which data are produced, gathered, pooled, analysed and interpreted, thus affecting the whole research cycle (Auffray et al., 2016). In what follows, we argue that the risks of testimonial and hermeneutical injustice in data practices can be mitigated by attention to methodological data fairness from the outset of a research project. To tackle these forms of injustice, fairness needs to be conceptualized in its relation to the goals, contexts and stakeholders involved in the data creation and use. More specifically, testimonial injustice can be mitigated by *data practices that actively counter existing prejudice about what counts as appropriate/relevant/adequate evidence*; while hermeneutical injustice can be mitigated by *data practices that leverage diverse sources of knowledge to counter existing prejudice due to ignorance or lack of understanding*.

Implementing methodological data fairness means considering how legal and social concerns can be addressed within each stage of research, thus helping

to ensure that potential gains are realised. This is critical to avoid unethical data uses as well as methodological problems in data analysis, especially since legal and institutional governance mechanisms struggle to keep up with rapidly developing methodologies and research findings (Halford and Savage, 2017), including potential harms arising from integrating different data sources (Leonelli, 2016). This does not guarantee the ethical use of data after a research cycle is complete, as researchers have limited control on how their work will be interpreted and re-used by others. However, attention to methodological data fairness can affect the infrastructures, expectations and commitments that accompany the data as they journey on, thereby helping to prevent immediate concerns and contributing towards longer term fairness. In this sense, methodological data fairness is an important component of scientific inquiry; as we show in the next section, asking what makes data use fair within everyday research practice is often equivalent to asking what makes data reliable, trustworthy and accurate.

## Methodological data fairness at each stage of the research process

### *Planning and design of research*

The first stage in achieving fair social media research relates to the fundamentals of the research; what is being asked, why, in what ways and for what purposes. Clarity in research goals informs robust design and data choices. In turn, the choice of data can affect and sometimes shift the goals of research (boyd and Crawford, 2012), a possibility that needs to be taken into account in the design and subsequent implementation of research – not least as part of Data Management Plans (DMPs).

Of further concern is evaluating the conceptual assumptions being made. It has been suggested that the apparent uncritical and a-theoretical nature of much big data and social media research, as well as a lack of self-reflexivity by researchers, has limited the development of the field and the utility of research outcomes (Tinati et al., 2014). Whether or not this is the case, we wish to problematise the widespread tendency among researchers to confuse the lack of well-established theories with the lack of conceptual assumptions. The lack of theory is understandable in a domain that is often data-driven and viewed as a precursor to the development of full-blown theories, but it does not mean that researchers do not make significant conceptual assumptions when developing their empirical research – such as the study of rural populations being irrelevant to the understanding of certain diseases, Twitter data being useful to identify

the emergence of symptoms in a population, or social media posts being reliable. Those assumptions need to be identified and evaluated when designing and carrying out a study.

Although new sources of data and methodologies are continually being developed and used, established scientific practices must continue to apply (boyd and Crawford, 2012). The intentions of the study, the assumptions being made about the phenomenon in question, and who may be affected or involved, should be explicitly considered throughout the research and reported in the final stages (Floridi, 2011). These questions may not always be clearly answerable, but posing them can help researchers anticipate the potential for harm or dis-benefit (Zook et al., 2017). This reflexivity should extend to the scientific context of researchers. The epistemic culture and commitments already in place in the research environment may need to be challenged to ensure that practices are lawful and fair.

Working with stakeholder and reference groups and, where appropriate, with diverse publics, can help refine research plans and surface unknown unknowns. For instance, it may be useful to poll a group of Twitter users to determine whether they are likely to make accurate tweets, or whether they tend to post ‘fake facts’ as a joke or for other purposes. Such a consultation may inform the extent to which researchers rely on Twitter content, how they structure data mining, and whether they use other data sources to triangulate findings. This also applies to the secondary use of data and to the application of any tools developed. Indeed, few tweets have geo-referenced data and yet in our past research, we have extrapolated from those few to all tweets in order to link them with environmental data, without reflecting on how representative the people with geo-referenced data may be of all Twitter users. This turned out to limit the re-usability and reliability of our results in the longer term.

In academic and some institutional settings, the use of SMD for research presents challenges to the established ethical review systems. There is much variation as to whether or not ethical review is sought; and even where established practices exist, approaches and standards are contested (Metcalf and Crawford, 2016; Vitak et al., 2016). Challenges include the implications of who or what is considered to be the research ‘participant’ and the ‘distance’ of researchers from data providers; a broadening of what is understood to be data; and the balancing of potential gains against uncertain harms (Neuhaus and Webmoor, 2012).

The highly dynamic nature of the social media industry poses further challenges to maintaining institutional oversight of fairness. Platforms change their Application Programming Interface (API) with little

or no notice, and what was once an acceptable and widely used data source can suddenly become prohibited. Changes are often not communicated widely to data users, except in the instrumental and technical aspects of how (e.g.) computer programs used to access the data must be altered to continue functioning; changes in the ‘back end’ functioning of the API that do not affect the direct interface may go unnoticed. Social media corporations protect commercially valuable information (such as algorithms used to control access and presentation of data, as well as complete raw datasets). Ethical oversight frequently operates in the absence of full information about the data and how they have arisen.

This has led to calls for an ‘agile’ ethics (Neuhaus and Webmoor, 2012) and questions as to whether researchers should rely solely on committees for their ‘ethical compass’ (Henderson et al., 2013). Promising novel avenues for oversight focus on ‘ethics ecosystems’, the sharing of information between ethics boards, regular reviews of the research process (Leonelli, 2016; Samuel et al., 2018), and more explicit consideration of fairness (Taylor and Pagliari, 2017). Ethics and data governance within research institutions need to capitalise on wider expertise to support methodological data fairness.

### *Data choices, sampling and acquisition*

Key considerations in undertaking fair social media research relate to data acquisition. We already stressed that Facebook, Instagram, WhatsApp, SnapChat and other major platforms are not available for research due to privacy constraints and lack of access (Fielsler et al., 2020). This creates a major distortion to what – and who – can be studied using SMD, which needs to be taken into account in research design.

The fact that some forms of SMD are available does not mean that they are *suitable*, and users must consider the strengths and limitations of the data. For instance, the population of interest should be defined theoretically rather than ‘solely by reference to technical capacities e.g. hashtag or location’ (Tinati et al., 2014). One aspect of assessing the suitability of data relates to the context of data production. Social media posts are not single, standalone comments or necessarily ‘truthful’ information. Their generation, and any responses they provoke, is part of a complex social process, which service providers play a key role in shaping (Taddeo and Floridi, 2016). Platforms are ‘black boxes’ that can change invisibly and arbitrarily for commercial benefit of the platform or its advertisers. Common approaches to gathering and initial processing of SMD strip the post or content from its context, potentially limiting understanding of its

meaning. Clarity in reporting data choices and sampling procedures is thus crucial to judging the fairness of social media research and application (boyd and Crawford, 2012).

We have surprisingly little information on the demographics of social media use at a population level, a situation compounded by the rapid evolution of usage patterns. Demographics may change by topic, geography and time of day within the same platform; for example, English-language posts on Twitter may come from US, UK or Australian citizens depending on when people are awake. Concerns have been raised about the trustworthiness of information sources (Floridi, 2011), with important implications for the research findings thereby derived. The authenticity of social media users – especially given the abundance of fake, pseudonymous and automated ‘bot’ accounts – is difficult to verify even with specialised identification tools (Hunter et al., 2018). Depending on the analysis performed and the intended purpose of research, these accounts may be seen as extraneous noise or as an accepted part of the social media milieu (Hunter et al., 2018). Attention to hermeneutical injustice demands that efforts are made to understand the population from which SMD are sampled, especially when the analysis aims to understand the beliefs, actions and intentions of socio-demographic groups and the results are generalised and used to guide decision-making.

The motivations of social media users cannot be verified without in-depth qualitative research. Except for specific citizen science initiatives (Tempini, 2017), few users post information intended to inform research. Some researchers take user consent to the Terms of Service (ToS) of the relevant platforms as a substitute for consent to data use, and exercise great care in staying within the boundaries determined by such a ToS. For instance, during the social sensing project, we stopped using Instagram data, despite their potential usefulness, since ToS stated that data should not be harvested at scale using automated methods. Although we could have gotten the data using a program that simulated many manual searches, we concluded that this might contravene the reasonable expectations of the user.

This is not a universal position regarding the ethics of restrictive ToS, with some researchers questioning the underlying power dynamics between user, platform and research (Fiesler et al., 2020). Moreover, both social media users and researchers are uncertain about what norms should apply to the use of their data (Nissenbaum, 2009; Zimmer, 2010). It cannot be assumed that social media users have read, understood or acted upon a ToS (Vitak et al., 2018); and consent from the original poster may not cover information acquired on their social networks due to the inherently

linked nature of social media material (Hunter et al., 2018). Similar issues have emerged in relation to bioethics (Anderson, 2015).

Different attitudes to SMD mining emerge from evaluations of whether practices were fair or not, whether or not the data are perceived to be ‘private’ or ‘public’, and in relation to transparency of the research process (Kennedy, 2016; Kennedy et al., 2017). The very retrieval of data via an API is ‘shaped by the methods that researchers use to access data, the economics and practicalities for the companies in sharing data, with whom and on what basis, both shaped by legal and sometimes even ethical considerations’ (Halford et al., 2017). Opacity in the storage, display and data provision/distribution protocols of social media companies limits understanding of what SMD represent. This lack of clarity is likely to introduce hidden biases (González-Bailón et al., 2014; Morstatter et al., 2013). For example, when geolocating data extracted from applications such as Twitter, it is unclear how the location is set and measured; the Twitter platform allows precise location information from smartphone GPS to be attached to individual tweets, but in the absence of GPS may utilise more general location information from the user profile. This can generate inequities in sampling and representation, depending on the scale and level of resolution of the location categories used (e.g. neighbourhood vs. street vs. household). As long as little information is provided on how these algorithms work for specific social media, it is difficult to evaluate what assumptions are being made and how they may affect the research undertaken.

### *Data processing and analysis*

Recognition of the potential for data processing to result in discriminatory outcomes has prompted greater consideration of the need to ensure data fairness (Berendt and Preibusch, 2017). Social media produces enormous quantities of data, much of which will be irrelevant to the research underway; this extraneous material must be filtered, or ‘cleaned’ (Kavanaugh et al., 2012). Computational methodologies, such as machine learning algorithms, can discriminate data according to predictive characteristics. However, machine learning is subject to similar sources of bias as other forms of data processing (Horvitz and Mulligan, 2015; Veale and Binns, 2017).

Supervised ML begins with a human-curated training dataset, from which the machine learns to detect patterns of interest. Thus, it can amplify or distort any biases used by human curators to produce the training data. Unsupervised ML finds statistical patterns in the data without much human involvement – but its

outputs are still as good as the data sample to which it is applied. Relatedly, many ML algorithms are inherently difficult to explain even if code is available for study, since decisions are decentralised and there is no clear IF-THEN logic (Munoz et al, 2016); and while in academic work code-sharing is common, in commercial uses of machine learning, the code is often proprietary.

The extrapolation and identification of individual characteristics (e.g. political affiliations) through the modelling of SMD are becoming increasingly accurate (Kosinski et al., 2013), with the linking of data from separate social media sources further enhancing these predictions (Leavey, 2013) and related infringement of privacy laws (Horvitz and Mulligan, 2015). Machine learning can also reveal information the social media users had not consciously shared. Codes of good practice regarding data processing and use, including from burgeoning research fields such as Discrimination-aware Data Mining (DADM) and Fairness, Accountability and Transparency in Machine Learning (FAT ML), are being developed (Berendt and Preibusch, 2017) but remain challenging to apply given ‘poorly designed matching system, decision-making systems that assume correlation necessarily implies causation, and data sets that lack information’ (Munoz et al., 2016). As noted, the processing of SMD should avoid stripping data of crucial contextual information.

### *Data storage, stewardship and management*

The creators of the FAIR principles emphasised that ‘good data management is not a goal in itself, but rather is the key conduit leading to knowledge discovery and innovation’ (Wilkinson et al., 2016). Good data stewardship is crucial in safeguarding public trust and maintaining fairness in social media research, and yet what constitutes good practice in the management of new forms of big data, including social media, is not well defined or acted upon (Auffray et al., 2016; Kinder-Kurlanda et al., 2017). Ostensibly, SMD should be treated as other potentially sensitive datasets with the interests of the ‘participant’ protected, including friends or family mentioned or implicated. The data and associated meta-data should be securely stored in line with predesigned DMPs. As we noted however, SMD raise specific challenges with respect to basic tenets of ethical data management such as consent and confidentiality.

To mitigate these challenges, bespoke DMPs and systems architecture must be designed towards secure data storage and retrieval. Kinder-Kurlanda et al. (2017), for example, developed robust methodologies to archive geotagged Twitter data, while Collmann et al. (2016) designed DMP guides which prompt

researchers to reflect on the implications of data re-use – including how to manage the deletion of data in relation to the ‘right to be forgotten’ of GDPR. A related challenge is posed by the restrictions constraining the sharing of SMD beyond the original posting. Twitter for instance allows researchers to download and store tweets via their API, but not to pass the tweets to anyone else. A common research practice is to share lists of tweet ID numbers and then query the Twitter API to retrieve the full tweet. This gives Twitter the ability to refuse to provide the tweet associated with an ID, which – while complying with GDPR regulations - makes it difficult for studies based on these data to be scrutinized and replicated.

### *Reporting and publishing*

This brings us to the issue of reporting. Although there are as yet no systematic studies of the reproducibility of social media research, assessments of similar fields concluded that it is ‘difficult if not impossible’ to independently reproduce studies from published papers (Olorisade et al., 2017). This likely applies to research using SMD which is published with ‘remarkably little methodological consideration of the data used’ (Halford et al., 2017). Lack of detail and transparency hampers researchers’ ability to interrogate the suitability of research components and veracity of the findings, hence limiting the fruitfulness of social media research and assessments of data fairness (Zimmer, 2010). The fields using SMD tend to perpetuate existing disciplinary habits for ethics and data governance. Researchers in public health or epidemiology are more sensitive to the demands of data fairness, while computer science has little history of engagement in this area. Nevertheless, much of the current social media research happens in computer science departments, with researchers doing social science research on public health despite not having research training in that area.

To improve methodological data fairness, researchers need to admit and report the limits of their understanding of the data at the point of publication. As noted, how data are created, curated, sampled and provided through APIs is often unclear, nor is there a good understanding of systemic biases in social media use. Institutions and venues in which social media research is undertaken should encourage their staff to recognise and investigate these issues. Another concern is the lack of recognised standards for assessing the quality of social media research (Cai and Zhu, 2015). The sheer variety of ‘methods, materials, goals, techniques used... as well as the diverse ways in which data can be evaluated depending on the goals of the investigation at hand’ present a significant challenge to performing peer review (Leonelli, 2017). Finding appropriate

**Table 1.** Steps towards achieving methodological data fairness in social media research.

Essential steps	<ul style="list-style-type: none"> <li>Compliance with legal data and privacy requirements</li> <li>Ongoing ethical oversight of the study including risk assessment of potential harms at regular intervals</li> <li>Appropriate choice and reporting of study design and methods</li> <li>Clarity in rationale for choosing data providers and reporting data sampling procedures</li> <li>Demonstrable understanding of populations from which data are sampled, and clarity in reporting limits in researchers' understanding of data at the point of publication</li> <li>Conformity with the Terms of Service of social media platforms</li> <li>Appropriate data management including preservation of dataset for later re-use</li> </ul>
Desirable steps	<ul style="list-style-type: none"> <li>Acquisition of explicit consent from data providers</li> <li>Maintenance of Data Management Plans (DMPs) with regular review</li> <li>Demonstrable understanding of how data are produced, including context (e.g. posts within wider conversations) and procedure (e.g. data curation within the platforms).</li> <li>Creation and training of interdisciplinary research teams including computer scientists, social scientists and health researchers</li> <li>Clarity of data processing algorithms and decisions at all stages</li> <li>Work with stakeholder or publics representative groups</li> <li>Replication of study in separate but contemporaneous samples</li> </ul>

ways to share SMD may help in this respect, while also lowering the digital divides in data access (Kinder-Kurlanda et al., 2017; Weller and Kinder-Kurlanda 2016). Making software (and related parameters) open source also improves the scrutiny of the research (Jiménez et al., 2017; Vasilevsky et al., 2017).

### **Conclusion: Data fairness fosters good research practice**

There is no doubt about the critical importance of the FAIR principles to effective data management. We have highlighted the equally critical importance of monitoring what happens once FAIR data mined from social media are being used to generate new knowledge; and particularly what the notion of fairness means in SMD mining. We argued that implementing methodological data fairness contributes to better and more responsible uses of SMD for health research, in ways that complement the FAIR framework. Methodological data fairness has a broader remit than the CARE principles, which focus strongly on ethics but restrict their scope to the datafication of indigenous knowledge. It also embraces data that are not necessarily filtered by large data repositories such as those targeted by the TRUST principles, though arguably all social media platforms could be viewed as large databases and should be administered accordingly. The fact that health research based on social media does not necessarily fall under the remit of the CARE and TRUST principles is one of many reasons why ethical concerns in this area are often overlooked by practitioners, especially within computer science and data science departments. Our focus on methodological data fairness aims to help remedy this situation,

by stressing the central role of ethics towards achieving reliable and innovative scientific findings.

At the start of our analysis, we highlighted two types of epistemic injustice that can affect data analysis and compromise results: testimonial and hermeneutical. Our discussion has shown that efforts to implement methodological data fairness at each stage of the research process, summarised in Table 1, help countering these forms of injustice. Testimonial injustice can be addressed by data practices that critically engage existing norms around what counts as appropriate, relevant or adequate evidence. Steps towards countering testimonial injustice include: considering whether platforms such as Twitter are a sufficient source of evidence for understanding population beliefs, preferences and behaviors; and adopting 'agonistic' machine learning, where 'companies or governments that base decisions on machine learning must explore and enable alternative ways of datafying and modelling the same event, person or action' (Hildebrandt, 2019). Hermeneutical injustice can be addressed by data practices that leverage diverse sources of knowledge to counter existing prejudice – for example, data collection and mining techniques that seek to mitigate existing digital divides and the privileging of specific socio-demographic, cultural or political stances within data sources.

We noted that data fairness needs to be promoted at a systems level by targeting a variety of stakeholders – in the case of social media research, the owners of online platforms through to users of research findings. The opacity of social media systems, from the implications of user interactions with the platforms to the nature of data gathered through APIs, is a serious challenge. Dialogue with social media users can help clarify

expectations and suitable processes for gaining consent. For researchers it is also essential to provide incentives like suitable funding calls, and publishing guidelines demanding rigorous ethical review. This creates a research culture where taking the time to consider data fairness becomes normal practice.

This is particularly relevant given the pressure that the COVID pandemic has placed on social media researchers working at the interface of public health. It has been suggested that researchers should relax their vigilance on data fairness, in the interest of speeding up the production of results that may help to respond to the emergency. The special circumstances linked to disasters have long been noted by data ethics analysts, with many suggesting that those undertaking big data research should know when to break rules (Zook et al., 2017: 7–8). We are uncomfortable with this suggestion. While the urgency of addressing a pandemic is all too obvious, the hasty production of unreliable results is not a useful scientific response (Leonelli 2021; O’Brian 2020). We argued that research outputs are more reliable, reproducible and easier to build upon when researchers pay attention to issues such as sampling, accurate reporting of the scope and limits of projects, and proper handling of consent – in other words, when researchers implement methodological data fairness. These efforts result in better science and more actionable evidence for decision-making.

We conclude that making research data fair as well as FAIR is inextricably linked to concerns around the adequacy of data practices. ‘Making data fair’ means critically identifying and regularly assessing the extent to which data practices are likely to participate in (or sometimes create) social stratification, clustering and surveillance of individuals or communities, and with what possible implications. This requires the implementation of processes of accountability, integrity, and justice as integral to the whole research process – not as an add-on, an institutional hoop to be jumped through or as a discrete stage in preparation or follow-up. Viewing fairness as part of the technical concerns underpinning data analysis is still far from common, as the exclusion of fairness from the FAIR principles exemplifies. Methodological data fairness helps to guarantee not only ethical safeguards, but also – and more fundamentally – the scientific soundness of research practices and outcomes. It is imperative that data scientists working with SMD engage actively in countering pernicious forms of data injustice that can severely damage the credibility and veracity of their knowledge claims.

### Acknowledgments

The authors gratefully acknowledge insightful feedback from three referees and the journal editors, which greatly helped to streamline the paper, and relevant discussions with research

partners including Asthma UK, MET office and Public Health England, as well as the Data Studies group at Egenis (especially Rachel Ankeny and Niccoló Tempini) and Kath Maguire and the Health and Environment Public Engagement (HEPE) group. The views expressed in the article are those of the authors, and not necessarily those of the partners.

### Declaration of conflicting interests

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

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the ESRC via grant ES/P011489/1, the European Research Council through grant award 335925 (The Epistemology of Data-Intensive Science), and the Alan Turing Institute (EPSRC grant EP/N510129/1).

### ORCID iDs

Sabina Leonelli  <https://orcid.org/0000-0002-7815-6609>

Lora Fleming  <https://orcid.org/0000-0003-1076-9967>

Hywel Williams  <https://orcid.org/0000-0002-5927-3367>

### Note

1. As there appears to be no commonly agreed upon terminology relating to social media in research, we define the language we use for clarity. The term *social media* relates to the process and content of information posted online. *Social media post* refers to any content created by human or bot published on *social media platforms* (defined as digital applications that allow users to share content) by *social media users* (individuals acting in a private capacity or on behalf of a group). The social media post is transformed into *social media data* (SMD) when it is gathered by a researcher. *Social media research* is the act and result of using SMD to better understand phenomena.

### References

- Anderson R (2015) *The collection, linking and use of data in biomedical research and health care: ethical issues*. London: Nuffield Council of Bioethics. Available at: <https://doi.org/10.17863/CAM.31760> (accessed April 2021).
- Auffray C, Balling R, Barroso I, et al. (2016) Making sense of big data in health research: Towards an EU action plan. *Genome Medicine* 8(1): 71.
- Berendt B and Preibusch S (2017) Toward accountable discrimination-aware data mining: The importance of keeping the human in the loop-and under the looking glass. *Big Data* 5(2): 135–152.
- Bicchieri C and Chavez A (2010) Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making* 23(2): 161–178.

- Boyd D and Crawford K (2012) Critical questions for big data. *Information, Communication and Society* 15(5): 662–679.
- Cai L and Zhu Y (2015) The challenges of data quality and data quality assessment in the big data era. *Data Science Journal* 14: 2.
- Clifford D and Ausloos J (2018) Data protection and the role of fairness. *Yearbook of European Law* 37: 130–187.
- Collmann J, FitzGerald KT, Wu S, et al. (2016) Data management plans, institutional review boards, and the ethical management of big data about human subjects. In Collmann J and Matei SA (eds) *Ethical Reasoning in Big Data: An Exploratory Analysis*. Cham: Springer International Publishing, pp. 141–184.
- Conway M (2014) Ethical issues in using twitter for public health surveillance and research: Developing a taxonomy of ethical concepts from the research literature. *Journal of Medical Internet Research* 16(12): e290.
- Cullity G (2008) Public goods and fairness. *Australasian Journal of Philosophy* 86(1): 1–21.
- D'Ignazio C and Klein LF. (2020) *Data Feminism*. Cambridge, MA: The MIT Press.
- Dijck Van J and Poell T (2016) Understanding the promises and premises of online health platforms. *Big Data & Society* 1–11. doi: 10.1177/2053951716654173
- Dijck van J, Poell T and De Waal M (2018) *The Platform Society: Public Values in a Connective World*. New York, NY: Oxford University Press.
- Dove ES, Townend D, Meslin EM, et al. (2016) Ethics review for international data-intensive research. *Science* 351(6280): 1399–1400.
- Fiesler C, Beard N and Keegan BC (2020) No robots, spiders, or scrapers: Legal and ethical regulation of data collection methods in social media terms of service. In: *Proceedings of the international AAAI conference on web and social media* 26 (Vol. 14, pp. 187–196).
- Floridi L (2011) The new ethical responsibilities of internet service providers. *Philosophy and Technology* 24(4): 369–370.
- Floridi L and Taddeo M (2016) What is data ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374: 20160360.
- Fricker M (2009) *Epistemic Injustice: Power and the Ethics of Knowing*. New York: Oxford University Press.
- Galbraith KL (2017) Terms and conditions may apply (but have little to do with ethics). *The American Journal of Bioethics* 17(3): 21–22.
- GDPR website (2019) EU data protection rules. Available at: [https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules\\_en](https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en) (accessed June 2019).
- Gelinas L, Pierce R, Winkler S, et al. (2017) Using social media as a research recruitment tool: Ethical issues and recommendations. *The American Journal of Bioethics* 17(3): 3–14.
- Golder S, Ahmed S, Norman G, et al. (2017) Attitudes toward the ethics of research using social media: A systematic review. *Journal of Medical Internet Research* 19(6): e195.
- González-Bailón S, Wang N, Rivero A, et al. (2014) Assessing the bias in samples of large online networks. *Social Networks* 38: 16–27.
- Halford S and Savage M (2017) Speaking sociologically with big data: Symphonic social science and the future for big data research. *Sociology* 51(6): 1132–1148.
- Halford S, Weal M, Tinati R, et al. (2017). Digital data infrastructures: Interrogating the social media data pipeline. In: *AoIR 2016: The 17th annual conference of the association of internet researchers*. Berlin, Germany, October 2017.
- Hays R and Daker-White G (2015) The care data consensus? A qualitative analysis of opinions expressed on twitter. *BMC Public Health* 15(1): 838.
- Henderson M, Johnson NF and Auld G (2013) Silences of ethical practice: Dilemmas for researchers using social media. *Educational Research and Evaluation* 19(6): 546–560.
- High Level Expert Group European Science Cloud (2016) *Realising the Open Science Cloud: First Report and Recommendations of the Commission High Level Expert Group on the European Open Science Cloud*. Brussels: European Commission. Available at: <https://digital-strategy.ec.europa.eu/en/library/first-report-high-level-expert-group-european-open-science-cloud> (accessed April 2021).
- Hildebrandt M (2019) Privacy as protection of the incomputable self: From agnostic to agonistic machine learning. *Theoretical Inquiries in Law* 20(1): 83–121.
- Horvitz E and Mulligan D (2015) Data, privacy, and the greater good. *Science* 349(6245): 253–255.
- Hunter RF, Gough A, O'Kane N, et al. (2018) Ethical issues in social media research for public health. *American Journal of Public Health* 108(3): 343–348.
- Information Commissioners Office (2017) *The Guide to Data Protection* (Vol. 2.9.8). London: Information Commissioner Office Website. Available at: <https://ico.org.uk/for-organisations/guide-to-data-protection/> (accessed April 2021).
- Jiménez RC, Kuzak M, Alhamdoosh M, et al. (2017) Four simple recommendations to encourage best practices in research software [version 1; peer review: 3 approved]. *F1000Research* 6: 876. doi: 10.12688/f1000research.11407.1.
- Kavanaugh AL, Fox EA, Sheetz SD, et al. (2012) Social media use by government: From the routine to the critical. *Government Information Quarterly* 29(4): 480–491.
- Kennedy H (2016) *Fair Game? User Evaluations of Social Media Data Mining Post, Mine, Repeat: Social Media Data Mining Becomes Ordinary*. London, UK: Palgrave Macmillan.
- Kennedy H, Elgesem D and Miguel C (2017) On fairness. *Convergence: The International Journal of Research into New Media Technologies* 23(3): 270–288.
- Kinder-Kurlanda K, Weller K, Zenk-Möltgen W, et al. (2017) Archiving information from geotagged tweets to promote reproducibility and comparability in social media research. *Big Data & Society* 4(2): 2053951717736336.

- Kosinski M, Stillwell D and Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America* 110(15): 5802–5805.
- Leavey J (2013) *Social Media and Public Policy: What is the Evidence*. London: Alliance for Useful Evidence.
- Lee S (2017) Studying “friends”: The ethics of using social media as research platforms. *The American Journal of Bioethics* 17(3): 1–2.
- Leonelli S (2016) Locating ethics in data science: Responsibility and accountability in global and distributed knowledge production systems. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences* 374(2083).
- Leonelli S (2017) Global data quality assessment and the situated nature of “best” research practices in biology. *Data Science Journal* 16: 32.
- Leonelli S (2021) Data science in times of pan(dem)ic. *Harvard Data Science Review* 3(1). DOI: 10.1162/99608f92.fbb1bdd6.
- Lin D, Crabtree J, Dillo I, et al. (2020) The TRUST principles for digital repositories. *Scientific Data* 7(1): 144. DOI: 10.1038/s41597-020-0486-7
- Lupton D and Michael M (2017) “For me, the biggest benefit is being ahead of the game”: The use of social media in health work. *Social Media + Society* 3(2): 1.
- Metcalf J and Crawford K (2016) Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society* 3(1): 2053951716650211.
- Mikal J, Hurst S and Conway M (2016) Ethical issues in using twitter for population-level depression monitoring: A qualitative study. *BMC Medical Ethics* 17: 22.
- Mittelstadt BD and Floridi L (2016) The ethics of big data: Current and foreseeable issues in biomedical contexts. *Science and Engineering Ethics* 22(2): 303–341.
- Mons B, Neylon C, Velterop J, et al. (2017) Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European open science cloud. *Information Services & Use* 37(1): 49–56.
- Morstatter F, Pfeffer J, Liu H, et al. (2013) Is the sample good enough? Comparing data from twitter’s streaming API with twitter’s firehose. In: *Seventh international AAAI conference on weblogs and social media*. Cambridge, MA, 8–11 July.
- Munoz C, Smith M and Patil D (2016) *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*. Executive Office of the President. Washington: The White House.
- Neuhaus F and Webmoor T (2012) Agile ethics for massified research and visualization. *Information, Communication and Society* 15(1): 43–65.
- Nissenbaum H (2009) *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford: Stanford University Press.
- Noble SF (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.
- O’Brian B (2020) Retractions and controversies over coronavirus research show that the process of science is working as it should. *The Conversation*. Available at: <https://theconversation.com/retractions-and-controversies-over-coronavirus-research-show-that-the-process-of-science-is-working-as-it-should-140326?utm> (accessed January 2021).
- O’Neill C (2017) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishing Group.
- O’Neill O (2004) Accountability, trust and informed consent in medical practice and research. *Clinical Medicine* 4(3): 269–276.
- Olorisade BK, Brereton P and Andras P (2017) Reproducibility of studies on text mining for citation screening in systematic reviews: Evaluation and checklist. *Journal of Biomedical Informatics* 73: 1–13.
- Pentzold C and Fischer C (2017) Framing big data: The discursive construction of a radio cell query in Germany. *Big Data & Society* 4(2): 205395171774589.
- Plantin J, Lagoze C, Edwards PN, et al. (2018) ‘Infrastructure studies meet platform studies in the age of google and facebook. *New Media & Society* 20 (1): 293–310.
- Picazo-Vela S, Gutiérrez-Martínez I and Luna-Reyes LF (2012) Understanding risks, benefits, and strategic alternatives of social media applications in the public sector. *Government Information Quarterly* 29(4): 504–511.
- Ryan A (2006) Fairness and philosophy. *Social Research* 73(2): 597–606.
- Samuel G, Ahmed W, Kara H, et al. (2018) Is it time to re-evaluate the ethics governance of social media research? *Journal of Empirical Research on Human Research Ethics* 13(4): 452–454.
- Srnicek N (2017) *Platform Capitalism*. London: Polity Press.
- Sinnenberg L, Buitendijk AM, Padrez K, et al. (2017) Twitter as a tool for health research: A systematic review. *American Journal of Public Health* 107(1): e1–e8.
- Taddeo M and Floridi L (2016) The debate on the moral responsibilities of online service providers. *Science and Engineering Ethics* 22(6): 1575–1603.
- Tamar S and Lucivero F (2019) Introduction to the special theme: The expansion of the health data ecosystem – rethinking data ethics and governance. *Big Data & Society* 6(2). doi: 10.1177/2053951719852969
- Taylor L (2017) What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society* 4(2): 2053951717736335. doi: 10.1177/2053951717736335
- Taylor J and Pagliari C (2017) Mining social media data: How are research sponsors and researchers addressing the ethical challenges. ? *Research Ethics* 14(2).
- Tempini N (2017) Till data do us part: Understanding data-based value creation in data-intensive infrastructures. *Information and Organization* 27(4): 191–210.
- Tinati R, Halford S, Carr L, et al. (2014) Big data: Methodological challenges and approaches for sociological analysis. *Sociology* 48(4): 663–681.
- Townsend L and Wallace C (2016) *Social Media Research: A Guide to Ethics*. Glasgow: University of Glasgow.
- Van Kleek M, Liccardi I, Binns R, et al. (2017) Better the devil you know: Exposing the data sharing practices of smartphone apps. In: *Proceedings of the 2017 CHI*

- conference on human factors in computing systems, 2 (pp. 5208-5220).
- Vasilevsky NA, Minnier J, Haendel MA, et al. (2017) Reproducible and reusable research: Are journal data sharing policies meeting the mark? *PeerJ* 5: e3208.
- Veale M and Binns R (2017) Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society* 4(2): 2053951717743530.
- Vitak J, Liao Y, Subramaniam M, et al. (2018) “I knew it was too good to be true”: The challenges economically disadvantaged internet users face in assessing trustworthiness, avoiding scams, and developing self-efficacy online. In: *Proceedings of the ACM on human-computer interaction*, 1; 2(CSCW): 1–25.
- Vitak J, Shilton K and Ashktorab Z (2016) Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, San Francisco, CA, pp. 941–953. DOI: 10.1145/2818048.2820078
- Weller K and Kinder-Kurlanda KE (2016) A manifesto for data sharing in social media research. In: *The Proceedings of the 8th ACM conference on web science*, 22–25th May, Hannover, Germany.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. (2016) The FAIR guiding principles for scientific data management and stewardship. *Sci Data* 3: 160018.
- Williams ML, Burnap P and Sloan L (2017) Towards an ethical framework for publishing twitter data in social research: Taking into account users’ views, online context and algorithmic estimation. *Sociology* 51(6): 1149–1168.
- Wolff J (2010) Fairness, respect and the egalitarian ethos revisited. *The Journal of Ethics* 14(3): 335–350.
- Zimmer M (2010) “But the data is already public”: On the ethics of research in facebook. *Ethics and Information Technology* 12(4): 313–325.
- Zook M, Barocas S, Boyd D, et al. (2017) Ten simple rules for responsible big data research. *Plos Computational Biology* 13(3): e1005399.
- Zuboff S (2019) *The Age of Surveillance Capitalism*. New York: PublicAffairs.