THEME 4: GOOD DATA ACTIVISM AND RESEARCH

Chapter Thirteen: Provocations for Social Media Research: Toward Good Data Ethics

Andrea Zeffiro

# Abstract

As academics continue to collect, scrape, integrate and analyse social media data, technical knowledge to acquire, analyse, secure and disseminate data is needed, but so too is a refined understanding of evolving ethical norms and values embedded within data protocols and practices. In fact, the requirement of technical understanding coupled with the contemporary rate of technological evolution itself presents complex ethical conundrums when it comes to the collection, maintenance, and dissemination of social media research data. Institutions and funding agencies champion research using social media data, but few, if any have ethical guidelines for social media research. In part, current policies are too broad for social media research and are therefore left open to interpretation. At the same time, however, codifying ethical considerations means disavowing a situational ethics principle, which recognizes how each social media research context is unique with a distinct set of traits and ethical challenges. This chapter outlines provocations for ethical decision making and provides prompts for researchers to engage with throughout the life-cycle of research.

# Introduction

This chapter sets forth considerations for social media research by identifying provocations for how academic research methods, practices, and norms can foster ‘good’ social media research data ethics. Over the last year I conducted a pilot study examining institutional research ethics guidelines from Canadian universities in order to assess the current trends, standards and norms for working with social media data in a national research context. My research has shown that institutions from funding bodies, to universities, to research ethics boards share a piecemeal approach to research ethics in the face of changing technologies. The considerations culminating from the pilot study register with international researchers, scholarly communities and institutions that are grappling with the same issues, and add to existing efforts at establishing ethical guidelines for research using social media platforms and data.[[1]](#footnote-1) My contribution to *Good Data* outlines provocations for ethical decision making and provides prompts for researchers to engage with when navigating ethical considerations during research design, but also throughout the life-cycle of research. These provocations and prompts do not aim to define a ‘one size fits all model’ for social media research, but rather intend to generate a dialoguebetween interdisciplinary researchers to better understand some of the key concerns pertaining to the integration of social media data into a range of scholarly projects, and engage with pressing questions as to how access to and use of social media data is mandated and governed, and how these practices impact scholarly research and the production of knowledge.

# ‘Big Data’

‘Big data’ remains an elusive term. My understanding of it emerges in part from a digital humanities research context, wherein emerging digital tools and techniques contribute to an expanding array of research methods and sources of evidence. ‘Big data’ in this sense concerns the application for research of computational tools, techniques and methods used in the extraction, analyses and visualization of data from social media platforms and the Internet. In this respect, ‘big data’ summons disciplinary domains like data science, data analytics, and data management, curation and stewardship, among others. But I also comprehend the term from a critical media and communication studies perspective, wherein it signals approaches that question and challenge the notion of data as a neutral phenomenon.[[2]](#footnote-2)

Outside of the academic contexts I describe, big data has been cited as an agent in healthcare, entertainment, education, personal wellness, and city planning to name but a few domains. The public safety sector, for instance, harnesses self-reported data from social media, publicly available data sets, physical sensors, and surveillance systems to respond quickly to emergency alerts. During Hurricane Harvey, when local 911 systems were failing, residents took to Facebook and Twitter to ask for assistance, and emergency respondents gathered crowd sourced information from social media platforms to refine their situational awareness and respond to requests for help.[[3]](#footnote-3) Thus, big data allegedly stands to improve nearly every facet of our lives.[[4]](#footnote-4) The general optimism for it has rendered it infallible, even in light of ‘public’ data breaches and leaks,[[5]](#footnote-5) focused inquiries into social media platforms taking liberties with user data,[[6]](#footnote-6) and disclosures of academic studies that have strained the limits of ethical conduct in the use of data.[[7]](#footnote-7) As of late, however, nebulous big data practices have received significant public attention and contempt.

In March 2018, *The Guardian* and *Observer* first reported on a Facebook and Cambridge Analytica data breach in which the political consulting firm was given access to the data of more than 87 million Facebook users.[[8]](#footnote-8) It was later revealed that data had been harvested through a personality app called *thisisyourdigitallife.* In 2013, Dr. Aleksandr Kogan developed the app separately from his research at Cambridge University, and through his company Global Science Research (GSR), and in collaboration with Cambridge Analytica, hundreds of thousands of Facebook users were paid to take the personality test and agreed to have their data collected for research purposes. The app worked under Facebook’s pre-2014 terms of service, which authorised the harvesting of data from not only the 270,000 users who installed it, but from their friends as well. Data mining was largely limited to what was published on public profiles, such as political beliefs, interests, and friends’ information. However, a small number of users had their data harvested from timelines, posts, and private messages. Cambridge Analytica was granted access to the data, and without authorization, the firm used it to build psychological profiles of voters in the lead up to the 2016 US Presidential Election and subsequently targeted individuals with personalized political advertisements. Under the pretences of ‘research’, corporate and political interests sustained the normalization of deceptive data gathering and marketing tactics, and Facebook, Cambridge Analytica and Kogan benefited financially from the exploitation of personal data. Facebook is a market leader in stockpiling personal data, which is at the core of its $40.6 billion annual business.[[9]](#footnote-9) On the heels of the scandal, Facebook’s Chief Technology Officer Mike Schroepfer revealed in a post published on the company’s news site that most of Facebook’s 2.2 billion members have had their personal data scraped by ‘malicious actors’ at some point.[[10]](#footnote-10) This concession further emphasizes the platform’s haphazard procedures for safeguarding user data as well as a wilful lack of disclosure to its participants about the extent to which personal data was collected and shared.[[11]](#footnote-11)

Public outrage in the aftermath of the scandal bolstered concerns scholars have voiced for numerous years in regards to ‘big data’.[[12]](#footnote-12) Interdisciplinary scholars invested in critical data studies consider the historical conditions that have contributed to ‘big data’, with focused attention on the public and private entities that exert levels of control over the production, analysis and management of data, and the ways in which emerging computational practices, techniques and methods contribute to forms of knowledge production in industry and the academy.[[13]](#footnote-13) In this sense, ‘big data’ as it is deployed in critical data studies frames urgent concerns about the production of knowledge in an effort to render transparent the ways in which data shape and are shaped by the instruments, practices, and contexts used to generate, collect, process, analyse, and store data. Orit Halpern refers to these assumptions about the value of data as effectuating ‘communicative objectivity’, which are new forms of observation, rationality and epistemology.[[14]](#footnote-14) For instance, it is common practice for the ‘digital traces’ or residues of information that are produced, abandoned or captured throughout social media exchanges to stand in for the digital identity of individuals and collectives.[[15]](#footnote-15) The risk in interpreting digital traces as the raw material of human identity,[[16]](#footnote-16) especially in academic research, is that it displaces the figure of the human subject and fundamentally reshapes practices and processes of knowledge production. Jacob Metcalf and Kate Crawford outline the ways in which ‘big data’ research challenge research ethics conventions in regard to the ‘human subject’.[[17]](#footnote-17) As the authors explain, the figure of the human subject in big data research is reconfigured into a ‘data subject’, marking a shift from an individual to a wider networked and distributed grouping or classification of people. ‘If the familiar human subject is largely invisible or irrelevant to data science’, ask Metcalf and Crawford, ‘how are we to devise new ethical parameters? Who is the ‘data subject’ in a large-scale data experiment, and what are they owed?’. [[18]](#footnote-18) Arguably, the ‘human subject’ has not been deemed inconsequential to data science, but rather digital platforms, networks and data have challenged conventional and static understanding of ‘research participant’ and in turn, created unprecedented tensions between the researcher and the researched.

# Research Ethics

In the decades following World War II, the development of principles of research ethics and the ethical treatment of persons were codified in national and international policies and documents, such as the Nuremberg Code (1948), the Declaration of Helsinki (1964), the Belmont Report (1979), and the UN Declaration of Human Rights (1948). These policies and documents, which were formulated in response to experiments performed on human test subjects illegally and without their knowledge or informed consent, sought to define the ethical and legal terms of research involving human subjects.[[19]](#footnote-19) In broad terms, ‘human research ethics’ encompass the norms and values that frame ethical considerations, such as ‘good’ behaviours, protocols and practices for research involving human subjects. In the context of academic research in particular, research ethics policy documents identify ethical issues in the design, coordination and management of research and signal practical and ethical considerations for responding to these issues.[[20]](#footnote-20) In Canadian universities, Research Ethics Boards (REBs) are responsible for reviewing research involving human participants and ensuring the safety and well-being of human participants. However, REBs in and of themselves are not perfect mechanisms to gauge ethical research. The long-standing model of evaluating all research through criteria designed for positivistic, biomedical modes of inquiry is deficient.[[21]](#footnote-21) ‘When research design and conduct is guided directly by regulatory bodies’ reflects Annette Markham, ‘issues of ethics can be obscured; ethics is more like directives than dilemmas or quandaries’.[[22]](#footnote-22) In turn, REBs function at times more like gatekeepers to the advancement of knowledge, rather than as institutional bodies assisting researchers to navigate ethical dilemmas.[[23]](#footnote-23)

In regard to social media research, the norms and values of ‘human research ethics’ upheld by REBs are strained by the complexity of interactions between individuals, networks and technical systems. For instance, any conventional understanding of ‘informed consent’ is circumvented by third-party disclaimers in platform policies and renders refusal of participation defunct.In turn, ethical standards are left to interpretation. For some, this may counteract concerns about ‘ethics creep’ and the continued bureaucratization of research.[[24]](#footnote-24) But at the same time, short of clear guidelines, certain forms of social media research are required to undergo REB review while others are not, which is not to say that all social media research should be exempt from REB review, but rather that such inconsistencies could very well denote exempted research as ethical simply by virtue of exemption. Additionally, a lack of guidance could encourage researchers to abide by a social media platform’s terms of service as ‘rules’ for research, yet these terms do not clarify the conditions for ethical research, but instead govern how a researcher is permitted to access and use the data.

In Canada, the principles to guide the design, conduct and review process of research involving humans are outlined in the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2),[[25]](#footnote-25) a joint policy of Canada’s three federal research agencies or Tri-Council, which is comprised of the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Social Sciences and Humanities Research Council of Canada (SSHRC). Canadian researchers and their institutions are required to implement the principles and articles of the TCPS 2 as a condition of federal funding.[[26]](#footnote-26) That said the Policy, which has been informed, in part, by leading international ethics norms and disciplinary and professional codes of conduct, is not merely a contract for funding. Rather, it serves as a benchmark for the ethical conduct of research involving humans. Ultimately, the TCPS 2 assists Canadian researchers and research institutions navigate the contours of ethical research, and it sets forth in unadorned terms a framework for research that emanates from three core principles: Respect for Persons, Concern for Welfare, and Justice.

‘Respect for Persons’ pertains to a person’s ability to exercise autonomy in a research context and encompasses, ‘the dual moral obligations’ to respect autonomy and safeguard individuals with developing, impaired or diminished autonomy. [[27]](#footnote-27) A crucial apparatus for respecting participants’ autonomy is to seek their free, informed and ongoing consent. In this regard, ethical conduct of research reflects a commitment to participation in research as a matter of informed choice that is based on a thorough understanding of the research project, including its potential benefits and risks to participants and others. The second principle, ‘Concern for Welfare’ relates to a ‘person’s experience of life in all its aspects’, and considers the impact on physical, mental and spiritual health, and physical, economic and social circumstances. [[28]](#footnote-28) Thus, concern for welfare means that researchers and research ethics boards not only protect the welfare of participants by providing them with sufficient information to adequately assess the risks and potential benefits associated with their participation in the research, but they also deliberate on the welfare of participants through the design, review and administration of research and in a manner that safeguards participants from any unnecessary or avoidable risks. [[29]](#footnote-29) Finally, the third principle, ‘Justice’ refers to a researcher’s commitment to treat participants fairly and equitably. Individuals or groups in vulnerable circumstances may need to be provided with focused attention in order to ensure just treatment. As the Policy insinuates, sustaining the principle of justice necessitates a consideration of vulnerability throughout the lifecycle of research.

By all measures, the TCPS 2 dispenses a model and guide for the ethical conduct of research involving humans.[[30]](#footnote-30) Thus, it serves as a template for researchers navigating the contours of established research norms. However, emerging computational tools, methods and sources of evidence such as social media data, strain our understanding of traditional research and ethics norms. A commitment to ‘Respect for Persons’ as it pertains to upholding participants’ autonomy by seeking their free, informed, and ongoing consent is weakened in those instances where consent is derived from vague privacy policies as outlined by social media service providers.[[31]](#footnote-31) For example, prior to public knowledge of Facebook’s Cambridge Analytica scandal, the company had already spearheaded studies that frayed the contours of ethical research. Most notably, Facebook’s dubious research practices became public in 2014 when it was revealed how researchers at Facebook and Cornell University manipulated the news feed of 689,003 subjects by deliberately changing the number of positive and negative posts in their News Feed over a week in January 2012.[[32]](#footnote-32) The study was intended to examine how the changes affected the emotional tone of users’ ensuing Facebook posts, in order to assess how emotion spreads through large populations.[[33]](#footnote-33) Some commentators were quick to note how ‘permission’ was derived from Facebook’s Data Use Policy, which outlined how the company used information they received for ‘internal operations, including troubleshooting, data analysis, testing, research and service improvement’.[[34]](#footnote-34) However, this version of the data policy, which explicitly references ‘testing’ and ‘research’ was updated in May 2012, four months after the study. Moreover, as Flick argues, even if Facebook’s terms of service had stipulated at the time of the study that participant data could be used for research, this would not have constituted valid informed consent because the study violated the normative expectations of users, and the terms of service do not allow for a participant to actively waive their expectations in any straightforward manner.[[35]](#footnote-35)

As gleaned from the ‘emotional contagion’ study, the capacity to uphold a ‘Concern for Welfare’ is diminished when a research participant is lured into a covert study; they are unable to assess the risks and potential benefits of their involvement and to willingly choose to participate or not. But even in overt research contexts it is increasingly difficult to notify research participants about foreseeable harms, especially when risk is measured by far less discernible outcomes, such as the erosion to information privacy.[[36]](#footnote-36) As made evident from Facebook’s Cambridge Analytica scandal, the possible harms extend far beyond immediate risks. The 270,00 users who installed Kogan’s app *thisisyourdigitallife* were unaware that data collected would be used for ‘research’ purposes, or that the app would harvest the data from their friends, or that data mined through Kogan’s company would be sold to Cambridge Analytica, who in turn would wield the data for controversial electoral purposes. One might define ‘immediate risk’ in this scenario as the harvesting of data from each of the friends of the 270,000 users who installed the app. However, if we understand ‘risk’, as Sheeva Sabati explains it, as extending ‘into the knowledge that is produced, disseminated, and enacted from the data, rather than merely what is collected’,[[37]](#footnote-37) then we have yet to fully grasp the risks or adequately assess the harms in their entirety.

As academics continue to collect, scrape, integrate and analyse social media data, technical knowledge to acquire, analyse, secure and disseminate data is needed, but so too is a refined understanding of evolving ethical norms and values embedded within data protocols and practices. In fact, the requirement of technical understanding coupled with the contemporary rate of technological evolution itself presents complex ethical conundrums when it comes to the collection, maintenance, and dissemination of social media research data. Institutions and funding agencies champion research using social media data, but few, if any have ethical guidelines for social media research. In part, current policies are too broad for social media research and are therefore left open to interpretation. At the same time, however, codifying ethical considerations means disavowing a situational ethics principle, which recognizes how each social media research context is unique with a distinct set of traits and ethical challenges. A standardized research ethics template cannot account for these unique characteristics or ethical considerations that arise on a contextual basis.[[38]](#footnote-38)

This dearth of guidance reflects broader trends in digital data policies and practices. As Sandra Soo-Jin Lee explains, the ‘vacuum in policy has placed unrealistic expectations on existing review structures to address the changing social and commercial arrangements that characterize these online platforms’.[[39]](#footnote-39) In turn, researchers are left struggling to understand their ethical obligations when it comes to the collection and management of ‘public’ data associated with social media. For example, big data collection and analysis of social media may reveal more information about people than what they choose to expose ‘publicly’. The interoperability of datasets demands ethical considerations beyond the matter of disclosing risk from a single dataset or individual control of personal data.[[40]](#footnote-40) Datasets that would otherwise be innocuous and adequately anonymized on their own can be used to reveal highly sensitive information when analysed in conjunction with other datasets. Thus, concerns over the privacy of personal data in large datasets depend not only on the safeguards applied to a primary dataset, but also those used in other auxiliary datasets. Likewise, permitting data to be identifiable beyond the context it was intended for and without explicit consent, such as integrating a screen grab or using a quote in scholarly dissemination from a social media user who is not a public figure, can expose the identity and profile of the user. Researchers are responsible for protecting the privacy and anonymity of unknowing participants, such as paraphrasing or narrativizing data reproduced for research output, and they ought to seek informed consent from each individual if data is used in ways that can identify them. This of course is further complicated by a platform that may insist on units of data being published only in their original form and attributed to its original poster.[[41]](#footnote-41) In these instances, researchers are tasked with safeguarding participants which may very well defy a platform’s definition of publicness.

# Research Context

My research has shown that few institutions in Canada have ethics guidelines that apply directly to social media research. Institutional research ethics documents that refer to digital data collection do so in terms of ‘internet research’ and redirect to the requirements of the TCPS 2. Increasingly more common in Canada are research data management (RDM) plans that outline protocols for data management and stewardship. In June 2016, the three federal research funding agencies – the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Social Sciences and Humanities Research Council of Canada (SSHRC) - released a *Statement of Principles on Digital Data Management*,[[42]](#footnote-42) in which it called on researchers, research communities, and institutions to develop data management plans and standards that ‘are consistent with ethical, legal and commercial obligations, as well as tri-agency requirements’.[[43]](#footnote-43) It is envisioned by the tri-agencies that these Principles will guide ‘the responsible collection, formatting, preservation and sharing of their data throughout the entire lifecycle of a research project and beyond’. [[44]](#footnote-44) The tri-agencies have since developed a draft *Research Data Management Policy,*[[45]](#footnote-45) and invited institutions, associations, organizations and individuals to offer feedback on it. This initiative is a significant development for RDM practices in Canada and internationally. However, data management and stewardship are not interchangeable with research ethics, but rather these practices ought to be integrated with ethical considerations of working with social media data and from the outset of research, that is, prior to data collection.

Again, the TCPS 2 puts forth clear recommendations and requirements that match established research norms for the ethical conduct of research involving humans. Social media research data, as I have tried to argue, challenge these conventions and norms, and researchers are obliged to interpret codes of ethical conduct that were written in the mid-20th century to guide the collection, analysis and representation of data in the 21st century.[[46]](#footnote-46) While institutions continue to refer to the TCPS 2, the contours of what constitutes ethical considerations for research involving social media data remains murky. For example, the Policy stipulates how, ‘information contained in publicly accessible material may, however, be subject to copyright and/or intellectual property rights protections or dissemination restrictions imposed by the legal entity controlling the information’.[[47]](#footnote-47) Social media platforms are the legal entities that control user generated content and set out and enforce the terms and conditions to data, including how scholars can use and access data for research. But scholarly access to data does not render data practices more transparent. On 11 April 2018, the second day of Mark Zuckerberg’s testimony before Congress, the *New York Times* published a piece in which journalist Brian Chen described in detail the process of downloading his Facebook data.[[48]](#footnote-48) Chen learned that approximately 500 advertisers had his contact information, which included his email address, phone number and full name, and he also discovered that Facebook had stored his entire iPhone’s address book. As Chen clarified, when he had installed the platform’s messaging app, rather than retain relevant information for only those contacts also on Facebook and Messenger, it kept a copy of the names and numbers of all 764 contacts, including his apartment door buzzer, a mechanic and a pizzeria. ‘My Facebook data’ reflects Chen, ‘also revealed how little the social network forgets’. [[49]](#footnote-49) We know very little as to what kind of participant data is collected and stored, how it is stored, for how long, the entities that gain access to the data, and the cost of access. Social media research implicates us into these deceptive practices.

The following section puts forward preliminary thoughts on two provocations for social media research data: negotiated relationships and transparency. These nascent considerations are guided by an expansive body of scholarship on digital ethics[[50]](#footnote-50) that considers the ethical issues in social media research[[51]](#footnote-51) and outlines practical considerations for determining the obligations researchers have towards their participants.[[52]](#footnote-52) These provocations join existing efforts to motivate research communities to consider their ethical obligations in light of the challenges social media research brings to research ethics norms and conventions.

## Negotiated Relationships

Navigating the ethical complexities of social media research occurs not in isolation, but in deliberation between three seemingly disparate relations: a researcher and their participants, a researcher and the data platform, and finally, a researcher and their Research Ethics Board (REB). Navigating these can be viewed also as a means through which a researcher refines their sense of accountability.[[53]](#footnote-53) Like negotiation, accountability is multi-directional. It requires rigorous thinking about the ramifications of the choices one makes in the lifecycle of research, rather than assuming that a platform’s terms and conditions or a research ethics board will fulfill the task of ensuring that research is conducted ethically.

## Researcher and Participants

As touched on previously, social media research displaces the figure of the human subject with data often standing in for the identity of individuals and groups.[[54]](#footnote-54) This distancing of the researcher from research participant makes it easier to approach social media data as non-human research, especially in circumstances when a researcher is working with social media corpora. One may recognize the corpus as produced intentionally or unintentionally by human participants throughout their social media exchanges, while also disavowing the place of the human subject in a traditional sense. As Michael Zimmer notes, ‘the perception of a human subject becomes diluted through increased technological mediation’.[[55]](#footnote-55) Researchers ought to consider how social media participants are conditioned into the role of data producers. Social media have been glorified for encouraging and promoting ‘sharing’, however, this goes only so far when one considers how sharing online is no longer a form of mere exchange but also a requisite for communication and distribution. ‘Sharing’, as Claire Birchall explains, ‘has to be understood today not as a conscious and conscientious act but as a key component of contemporary data subjectivity’.[[56]](#footnote-56) Birchall reframes sharing protocologically, as ‘the constitutive logic of the Internet’.[[57]](#footnote-57) Sharing, in other words, is a standard of the system. Activities and practices online that appear to be driven by a free will to share are in effect preconditions to participation and standardized practice. If social media data are generated in large part from individuals who are compelled to ‘share’ data as a prerequisite for participation, then how might this infringe on a researcher’s capacity to uphold ‘Respect for Persons’ and a ‘Concern for Welfare’? Recognition of one’s dataset as having been generated by human participants who are likely unaware of how their thoughts, emotions, and observations have been quantified, and in turn, applied by researchers, is imperative in advancing evolving ethical benchmarks.

## Researcher and Data Platform

The term ‘platform’ is synonymous with social media and is often used to refer to those web-based interfaces through which individuals are able interact with other people and share content. For instance, Twitter, Facebook, YouTube and Instagram are some of the more prominent social media platforms in North America. Thus, a social media platform has social characteristics, as I described, but it also has specific technical attributes. Technologically, a platform provides a mark-up language for creating applications, an Application Programming Interface (API) for third-party application integration ,[[58]](#footnote-58) and a back-end administrative console for managing site architecture and functionality. As scholars have argued, a ‘platform’ is not simply a social or technological tool. Rather, digital intermediaries employ the term as a discursive strategy to frame their services in a particular manner and present themselves as transparent entities in the facilitation of public dialogue and communication, rather than as entities who serve and profit from content providers, advertisers and others.[[59]](#footnote-59)

When researchers seek out social media data from a particular platform, they are in effect entering into a relationship with that platform. First and foremost, a researcher is governed by the terms and conditions set forth by the platform. This is beneficial in instances when agreements articulate how one is permitted to access and use data for research via the platform’s standards, but these do not necessarily align with ‘ethical research’. Platforms are not simply neutral data portals through which researchers are permitted access to troves of data. Platforms are data gatekeepers that create and specify constraints as to who can access data, in which forms, and under which conditions.[[60]](#footnote-60) As Taina Buchner argues, researchers employing data collection tools like APIs need to know how these tools collect and provide access to the data and functionality contained by platforms, but they also have a responsibility to understand how a seemingly neutral tool like an API is not a conduit for data, but is instead is a ‘technique for governing the relations they contain’.[[61]](#footnote-61) Following Tarleton Gillespie,[[62]](#footnote-62) platforms and their data tools have ‘politics’, meaning they can be understood as having ‘powerful consequences for the social activities that happen with them, and in the worlds imagined by them’.[[63]](#footnote-63) Thus, rather than asking what these data platforms are, researchers are better served to ask what these platforms do.

A researcher’s contractual obligation to a platform ought to be reframed as a partnership with a data gatekeeper, rather than as an agreement with terms and conditions. A researcher wanting to access data will have to do so according to the platform’s policies, which may misalign with ethical research. Researchers are therefore left to negotiate what Mark Andrejevic has called the ‘big data divide’.[[64]](#footnote-64) There exist unequal ways of accessing and using data that intensify power imbalances. Thus, a ‘big data divide’ describes the asymmetrical relationship between those who collect, mine, store and analyse data, and those whom data collection targets. Perhaps framing the relationship in this way will render transparent the ways in which social media research data are not neutral, objective, and pre-analytic in nature. Data are a by-product of the politics of platforms. What if research communities conceived of these platforms not simply as sources of research evidence, but as collaborators in the construction of emerging research practices and knowledge production? Would this compel researchers to dig deeper into the politics of platforms as a condition of working with social media data? Perhaps a first step in challenging prescriptive data regimes is for researchers to make concerted efforts to reflect on and make clear in their methodologies the key role platforms play in the co-construction of knowledge.

## Researcher and Research Ethics Boards

A researcher’s relationship with their Research Ethics Board (REB) ought to be positioned as a continuous dialogue, rather than as an obstacle to research. This is a tall order given how fraught this relationship can be. Indeed, the REB model itself is discordant at times with contemporary research practices and overburdened by risk management and bureaucratic box ticking.[[65]](#footnote-65) In many instances, REBs are also struggling to understand the ethical complexities of social media research, and uncertainty may lead to trepidation. In this respect, social media research may be deemed too risky because it is not well understood. Thus, with few guidelines or protocols for social media research specifically, researchers find themselves seeking expertise and guidance on ethical considerations. In part, I view this productively because it requires that researchers confront the challenges and conundrums of evolving research norms through practical application and beyond the limited scope of regulatory guidelines.[[66]](#footnote-66) Researchers have leeway in interpreting and applying existing ethics protocols to emerging research practices, permitting them to establish new benchmarks for research. At the same time, however, a lack of standardized practices[[67]](#footnote-67) with regards to social media research leads to inconsistent views about how to handle ethical issues,[[68]](#footnote-68) while interpretations of existing protocols for new research contexts may also betray broader ethical conventions, as evidenced by the emotional contagion study and Facebook’s Cambridge Analytica scandal.

What is needed are guidelines to allow for social media research to remain flexible but that would foreground ethical considerations to steer research design and methodological considerations, even in those instances when data is deemed ‘public’. Ethical considerations for and about social media research must trouble the public/private dichotomy instituted and governed by the terms and conditions established by platforms. In other words, simply because information is stipulated as ‘public’ does not absolve researchers of ethical concerns because of the presumed ‘publicness’ of data.[[69]](#footnote-69) For instance, according to the TCPS 2, REB review is ‘not required for research that relies exclusively on secondary use of anonymous information’, ‘so long as the process of data linkage or recording or dissemination of results does not generate identifiable information’.[[70]](#footnote-70) According to this provision, research using social media corpora, which falls within the parameters of ‘secondary use of anonymous information’, is exempt from REB review. However, if we reconsider how social media data is generated by human participants who are likely unaware of the parameters of secondary data, and how platforms are also data gatekeepers that co-produce knowledge, should we not then reexamine REB exemption? Or, should this fall onto researchers to advocate for ethical considerations, like REB review, that go beyond Tri-Council recommendations and requirements? As explored in this chapter, informed consent is deficient in social media research contexts. As agencies push for ‘open data’ as a requirement of funding and compel researchers to share research datasets, even if consent is obtained for a particular research study, how is it transferred when a data set is shared? Can it be transferred? In these emerging contexts, how can a researcher possibly guarantee confidentiality? The simple answer is that it cannot be guaranteed. Perhaps in the process of acquiring informed consent for social media research, a ‘no guarantee’ clause needs to be accentuated. Penn State for instance recommends the following statement be used: ‘Your confidentiality will be maintained to the degree permitted by the technology used. Specifically, no guarantees can be made regarding the interception of data sent via the Internet by any third parties’.[[71]](#footnote-71) I would go so far as to endorse a version of a ‘no guarantee’ clause on all research dissemination.

# Transparency

‘Are you willing to change your business model in the interest of protecting individual privacy?’

– Democratic Representative Anna Eshoo[[72]](#footnote-72)

‘Congresswoman, I'm not sure what that means.’

– Mark Zuckerberg[[73]](#footnote-73)

Privacy in relation to social media data has received significant attention as of late.[[74]](#footnote-74) The Facebook and Cambridge Analytica revelations have attracted scrutiny over the lack of autonomy over one’s data and (re)focused debates about privacy with demands for formal governance and regulation. But an emphasis on privacy alone is limiting in our ability to rethink not only our relationship to the data we generate, but also the processes and tools through which we access social media research data. For this reason, I am more invested in the concept of transparency. Transparency discloses the parameters of privacy but also the ways in which social media data operate as a kind of currency, that is, as an accepted source of evidence in academic research, and as a medium of exchange.

Privacy, as it is guaranteed by social media platforms, at least in theory, tends to register as an assurance that data is secure from ‘malicious actors’, and that it is collected, shared and used in ways we have consented to. And yet, platforms alter their terms of service and renegotiate the conditions of their user agreements to work in their favour,[[75]](#footnote-75) and they grant a multitude of unfamiliar entities access to our data including researchers.[[76]](#footnote-76) In academic research contexts, researchers have an ethical duty of confidentially to participants, which includes upholding a research participant’s right to privacy and safeguarding their information. This version of privacy is at odds with the conditions supported by data platforms. Thus, if researchers integrate data from platforms that overstep moral imperatives like privacy, then are researchers also straining the ethical contours of privacy norms and conventions in academic research?

What is ‘privacy’ in the context of social media research? If we follow Helen Nissenbaum’s lead, concerns over privacy are not simply concerns about control over personal information. A ‘right to privacy’, reflects Nissenbaum, ‘is neither a right to secrecy nor a right to control, but a right to the appropriate flow of personal information’.[[77]](#footnote-77) Nissenbaum advocates for a ‘contextual integrity approach’ to privacy, wherein ‘we locate contexts, explicate entrenched informational norms, identify disruptive flows, and evaluate these flows against norms based on general ethical and political principles as well as context specific purposes and values’.[[78]](#footnote-78) In this respect, privacy is tied to the norms governing distinct social contexts, but at the same time, privacy online is not something distinct from privacy offline. Rather, social norms, including informational norms, are inextricably linked with existing structures of social life.[[79]](#footnote-79) If information is flowing beyond the context it was intended for and without regard for a context’s particular norms and values, then privacy is not upheld or safeguarded.[[80]](#footnote-80)

Revisiting momentarily the question Democratic Representative Anna Eshoo asked Mark Zuckerberg during his Congressional hearing in April 2018, and reframing it for an academic context, we might find ourselves inquiring, ‘are you willing to change your research model in the interest of protecting individual privacy?’. If Eshoo’s query to Zuckerberg insinuates that Facebook’s business model is at odds with safeguarding privacy, then that same question reformulated for an academic context implies those same business models infringe on the established norms and conventions of privacy in academic research. Zuckerberg’s reply, ‘I’m not sure what that means’, deflects accountability. Researchers are not at liberty to divert privacy concerns. One possible means of respecting privacy in social media research is to approach it contextually, meaning that a researcher’s reading of ‘expectations of privacy’ is a negotiation between a particular platform’s terms of privacy, its audience, and aims. And by a nuanced understanding of how privacy expectations vary from platform to platform, group to group, and individual to individual. Despite a researcher’s best efforts to uphold expectations of privacy, its limits are tested by virtue of a researcher’s negotiated relationship with data platforms. I am putting forward transparency as a conceptual counterpoint to work through the limitations of privacy guarantees.

Transparency in scholarly research is often conflated with ‘openness’ in the sense of open source, open access and open data, and with the replication of results. To be exact, I am employing transparency here as a marker of intentionality on the part of the researcher, but also in terms of the platform and its policies and practices, which may not be transparent to the researcher. When a researcher deliberately carries out research in a way for others to comprehend the actions, negotiations, and deliberations performed as part of the research process, they are in effect enacting transparency. How then does one sustain transparency as an ethical consideration? At the very least, a researcher considers a process-oriented approach to research, wherein the process itself is just as important as the final output. By this I mean that one’s research process is fore grounded, particularly in scholarly output and dissemination. Researchers make clear their methods of data collection and analysis, and reflect on the negotiations between key actors and diverse relations facilitating research and the co-production of knowledge. Transparency in this respect helps to deconstruct processes of knowledge production: how knowledge is produced, who produced it, and for whom. Rather than sustaining ‘communicative objectivity’,[[81]](#footnote-81) transparency discloses new modes of observation engendered by data tools and sources through which scholarly communities are observing and analysing the world.

Transparency describes how researchers engage in ongoing processes of reflexive practice and revision by foregrounding research intentions, limitations, negotiations, and methods of data collection and analysis. I stand by the term as a provocation, but I also seek to trouble it. As I argued, transparency in social media research is radically important because it is a characteristic lacking from social media platforms. Platforms, as previously discussed, tend to be obtuse technical systems that purport to facilitate social engagement without full disclosure as to how participation is mediated for other individuals and entities that profit from data production. For instance, users of social media are well aware of how in exchange for a ‘free’ service like Facebook, the company collects their data and uses it to serve them ads both on Facebook and around the web. It is a seemingly simple exchange. But social media platforms like Facebook have proven to be poor stewards of user data, often demanding and doing more with it but without unambiguously disclosing their practices. When Facebook supposedly discovered in 2015 that Aleksandr Kogan provided data to Cambridge Analytica, it took until March 2018, after the publication of stories from *The Guardian*, *Observer* and *The New York Times*, for Facebook to both disclose it and suspend Kogan and Cambridge Analytica from its platform. Arguably, there is a fundamental lack of transparency.[[82]](#footnote-82) But at the same time, Facebook increasingly touts transparency as a catchphrase to signal to users that it is committed to disclosing its practices and that its activities are open to public scrutiny.[[83]](#footnote-83) Facebook recently released ad transparency tools that enable users to see more information than ever before about how advertisers are using the platform.[[84]](#footnote-84) For the average user, these tools may reveal the amount of advertising activity carried out on these platforms, but they do not make transparent exactly how ads operate on the platform. Moreover, just because users are given access to more information does not mean it is easy to parse. In turn, Facebook’s transparency serves to uphold its core policies and practices without revealing any more about how our data is trafficked.

To reiterate, transparency as an ethical consideration in social media research is radically important, but because it has been co-opted by technology companies, perhaps we need an additional term to address the messiness and complexities of working with social media data. To this end, I propose ‘c/overt research’. In their reflection on their experiences in a c/overt research project, Virtová, Stöckelová, and Krásná conceived of the term as a way to interrogate how IRB standards and the ‘ethical fiction’ of informed consent serve to insulate researchers from having to openly acknowledge uncertainties in field work.[[85]](#footnote-85) Thus, ‘c/overt research’ troubles the distinctions between overt and covert forms of research and insinuates that all research is covert in some ways, becoming overt only during the research process itself.

C/overt research as I adopt it fractures the myth that researchers are absolved of ethical concerns by virtue of seeking REB approval and abiding by ethical guidelines or codes of ethic. In following Alexis Shotwell’s work on purity politics, we are better served to view the aspiration for ethical purity as simultaneously, inadequate, impossible and politically dangerous.[[86]](#footnote-86) As I have argued, researchers working with social media data enter into a relationship with a platform. Rather than view the REB process as a means through which one is able to neutralize this relationship, we might consider highlighting the ways in which relying on the terms set forth by a social media platform legitimizes their nebulous data practices, and how this renders us complicit in these practices. As Shotwell explains, ‘[s]ince it is not possible to avoid complicity, we do better to start from an assumption that everyone is implicated in situations we (at least in some way) repudiate’. [[87]](#footnote-87) Complicity, and indeed complexity, is not something we should avoid in research contexts. Understanding not only how researchers are complicit, but REBs and institutions as well, is a ‘starting point for action’. [[88]](#footnote-88) In this respect, the ways in which researchers conduct themselves in the c/overt practice of their research is a mode of achieving, rather than applying, ‘ethical research’.

# **Good Data**

‘Negotiated relationships’ and ‘transparency’ are but two provocations for social media research. These terms outline some of the ethical complexities of working with social media data and the ethical concerns researchers may consider when entangled within contemporary data practices. Yet neither identify a pathway to good data practices. This section formulates questions for researchers to navigate ethical considerations during research design, that is, prior to data collection, but also to spur reflexivity throughout the life-cycle of research. These prompts are meant as an exploratory guide towards establishing definitive good data ethics. ‘[E]thics, when practiced’, write Markham, Tiidenberg and Herman, ‘becomes a matter of method’.[[89]](#footnote-89) Good data ethics can engender good data methods and vice versa.

*Research Questions:* What are some of the questions driving the research? What conceptual and/or theoretical frameworks are shaping these questions? How have other disciplines explored similar questions and to what end?

*Research Data:* What are my data sources? How will I acquire them? Is REB approval required? If not, will I seek out approval? How will data be managed and by whom? Who will be responsible for anonymizing and encrypting data? How and where will data be stored? Who will have access to the data and in what form?

*Research Tools:* What computational tools and techniques will be employed for research? Why these and not others? What skills and expertise are required? Who will conduct this portion of the research and how will they be acknowledged? What are other ways of doing the research?

*Research Relations:* What are some of my negotiated relationships? To whom do I feel accountable towards? With whom do I share this accountability? Where am I in the research and what is my situated perspective?

*Research Participants:* Who and/or what constitute my research participants? Is REB approval required? If not, will I seek it out anyway? How will participants be made aware of their involvement in the research? If this is not practical, then how will participation be made c/overt? What do I feel is my duty to these participants? How will I safeguard contextual integrity?How will I uphold participant autonomy? What are some possible ways in which I may disappoint research participants?

*Research Beneficiaries:* For whom is this research for? Who and/or what is the driving force? Why do I care about it? How will it benefit me? How will it be of benefit to others? Who will derive advantage from it?

*Research Dissemination:* How do I intend to share results of research? In what forms and with whom? How will I uphold contextual integrity when sharing results? Will a ‘no guarantee’ accompany research dissemination?

# Conclusion

The provocations and prompts offered here are far from exhaustive, but rather are a preliminary effort at identifying some of the tensions inherent in upholding good data research practices. As I have discussed throughout this chapter, there is a lack of ethical guidelines for social media research. In turn, researchers are often dependent on a mixed bag approach when it comes to ethics. That said, even if codes of ethics for social media were institutionalized, the ethical conundrums addressed throughout this chapter are not simply solved by reference to ethics documents and policies. My hope is that researchers and their institutions approach social media research as iterative and deliberative, rather than cement data ethics protocols or a one-size-fits-all model. Flexibility of this kind will enable research communities to transparently respond to emerging data tools, instruments, practices, contexts and epistemologies and develop further strategies for good data ethics that will empower researchers to respond to the prescriptive data regimes set forth by social media platforms that indubitably impact scholarly research practices and the production of knowledge. Finally, instead of fixating on the deficit of guidance, perhaps we are better served to interpret these challenges as opportunities, and rather than focus on codes of conduct imposed from the outside, we focus on the hidden ethical practices from the inside, that is, through ethical practices as they unfold in social media research contexts.[[90]](#footnote-90) Indeed, in this way, ethics are achieved, not applied.

# References

Adams, Richard. ‘Cambridge University asks Facebook for Evidence About Role of Academic’, *The Guardian*, 20 March 2018, https://www.theguardian.com/uk- news/2018/mar/20/cambridge-university-asks-facebook-for-evidence-about-role-of- academic-alex-kogan.

Adams, Tim. ‘Facebook’s Week of Shame: The Cambridge Analytica Fallout’, *The Guardian*, 24 March 2018, https://www.theguardian.com/technology/2018/mar/24/facebook-week-of-shame- data-breach-observer-revelations-zuckerberg-silence.

Andrejevic, Mark, and Kelly Gates. ‘Big Data Surveillance: Introduction’, *Surveillance & Society* 12.2 (2014): 185-96, DOI: doi:10.24908/ss.v12i2.5242.  
Andrejevic, Mark. ‘Big Data, Big Questions: The Big Data Divide’, *International Journal of Communication* 8 (2014): 1673-1689.   
Australian National Data Service. *Data Sharing Considerations for Human Research Ethics Committees*, June 2018. http://www.ands.org.au/guides/data-sharing-considerations- for-hrecs.  
Barocas, Solon, and Helen Nissenbaum. ‘Big Data’s End Run Around Procedural Privacy Protections’, *Communications of the ACM* 57.11 (2014): 31-33.   
Barrett, Brian. ‘Facebook Owes You More Than This’, *Wired,* 19 March 2018, https://www.wired.com/story/facebook-privacy-transparency-cambridge-analytica/. Bell, Kirsten. ‘Resisting Commensurability: Against Informed Consent as an Anthropological Virtue’, *American Anthropologist* 116.3 (2014): 511-522.  
Bhattacharya, Kakali. ‘Consenting to the Consent Form: What are the Fixed and Fluid Understandings Between the Researcher and the Researched?’, *Qualitative Inquiry* 13.8 (2007): 1095-1115.  
Biddle, Sam. ‘Stop Using Unroll.me, Right Now. It Sold Your Data to Uber’, *The Intercept*, 24 April 2017, https://theintercept.com/2017/04/24/stop-using-unroll-me-right-now- it-sold-your-data-to-uber/.   
Birchall, Clare. ‘Shareveillance: Subjectivity between Open and Closed Data’, *Big Data & Society* 3.2 (2016).

Bloomberg Government. ‘Transcript of Zuckerberg’s Appearance Before House Committee’, *The Washington Post*, 11 April 2018, https://www.washingtonpost.com/news/the- switch/wp/2018/04/11/transcript-of-zuckerbergs-appearance-before-house- committee/?noredirect=on&utm\_term=.71d99a22271d.   
Boellstorff, Tom. ‘Making Big Data, in Theory’, First Monday 18.10 (2013).

Booth, Robert. ‘Facebook Reveals News Feed Experiment to Control Emotions’, *The Guardian*. 30 June 2014, https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions- news-feeds.   
boyd danah and Kate Crawford. ‘Critical Questions for Big Data: Provocations for a Cultural, Technological and Scholarly Phenomenon’, *Information, Communication & Society* 15.5 (2012): 662–679.  
boyd danah. ‘Where do we find ethics?’ *Points*, 5 April 2016, https:// points.datasociety.net/where-do-we-find-ethics-d0b9e8a7f4e6.

Broad, Ellen. ‘Australia, we need to talk about data ethics’, *The Ethics Centre*, 25 January 2017, http://www.ethics.org.au/on-ethics/blog/january-2017/australia-data-ethics

Bruns, Axel. ‘Faster than the Speed of Print: Reconciling ‘Big Data’ Social Media Analysis and Academic Scholarship’, *First Monday* 18.10 (2014).

Brunton, Finn, and Helen Nissenbaum. *Obfuscation: A User’s Guide for Privacy and Protest*, Cambridge MA: MIT Press, 2016.

Brunton, Finn, and Helen Nissenbaum. ‘Vernacular Resistance to Data Collection and Analysis: A Political Theory of Obfuscation’, ***First Monday*** 16.5 (2011).

Bryant, Randall E., Randy H Katz and Edward D. Lazowska. ‘Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science, and Society’, Computing Research Consortium, 22 December 2008, https://cra.org/ccc/wpcontent/uploads/sites/2/2015/05/Big\_Data.pdf.

Bucher, Taina. ‘Objects of Intense Feeling: The Case of the Twitter API’, *Computational Culture* 3 (November, 2013), http://computationalculture.net/objects-of-intense-feeling-the-case-of-the-twitter-api/.

Cadwalladr, Carole and Emma Graham-Harrison. ‘Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach’, *The Guardian*, 17 March 2018, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

Chen, Brian X. ‘I Downloaded the Information that Facebook has on Me. Yikes’, *The New York Times*, 11 April 2018, https://www.nytimes.com/2018/04/11/technology/personaltech/i-downloaded-the-information-that-facebook-has-on-me-yikes.html.

Clark, Karin, Matt Duckham, Marilys Guillemin, Assunta Hunter, Jodie McVernon, Christine O’Keefe, Cathy Pitkin, Steven Prawer, Richard Sinnott, Deborah Warr and Jenny Waycott. *Guidelines for the Ethical Use of Digital Data in Human Research*, The University of Melbourne, 2015, https://www.carltonconnect.com.au/read-ethical-use of-digital-data-in-human-research/.

Collins, Keith and Larry Buchanan. ‘How Facebook Lets Brands and Politicians Target You’, *The New York Times*, 11 April 2018, https://www.nytimes.com/interactive/2018/04/11/technology/facebook-sells-ads-life-details.html.

Crawford, Kate, Kate Miltner and Mary L. Gray. ‘Critiquing Big Data: Politics, Ethics, Epistemology’, *International Journal of Communication* 8(2014): 1663–1672.

Dalton, Craig, and Jim Thatcher. ‘What Does a Critical Data Studies Look Like, and Why Do We Care? Seven Points for a Critical Approach to “Big Data”’, *Society & Space Open Site*, (2014), http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/.

Dalton Craig, Linnet Taylor and Jim Thatcher. ‘Critical Data Studies: A Dialog on Data and Space’, *Big Data & Society* 3.1 (2014): 1–9.

Das, Sauvik and Adam Kramer. ‘Self-censorship on Facebook’, *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, 2013, pp. 120-127, https://research.fb.com/publications/self-censorship-on-facebook/.

Davies, Huw et al. ‘Ethics Guidelines and Collated Resources for Digital Research’, British Sociological Association, 2017, https://www.britsoc.co.uk/media/24309/bsa\_statement\_of\_ethical\_practice\_annexe.pd f   
Dingwall, Robert. ‘The Ethical Case Against Ethical Regulation in Humanities and Social Science Research’, *Twenty-First Century Society* 3.1 (2008): 1-12.

Egan, Erin. ‘Enhancing Transparency in Our Data Use Policy’, *Facebook Newsroom*, 11 May 2012, https://newsroom.fb.com/news/2012/05/enhancing-transparency-in-our-data- use-policy/.   
Ess, Charles. *Digital Media Ethics*, Cambridge: Polity Press, 2014.  
Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York: St. Martin’s Press, 2018.  
Feeley, Malcom M. ‘Legality, social research, and the challenge of institutional review boards’, *Law & Society Review* 41 (2007): 757-776.  
Ferguson, Andrew Guthrie. *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*, New York: NYU Press, 2017.  
Flick, Catherine. ‘Informed Consent and the Facebook Emotional Manipulation Study’, *Research Ethics* 12.1 (2016): 14-28.  
Frenkel, Sheera, Matthew Rosenberg and Nicholas Confessore. ‘Facebook Data Collected by Quiz App Included Private Messages’, *The New York Times*, 10 April 2018. https://www.nytimes.com/2018/04/10/technology/facebook-cambridge-analytica- privatemessages.html?action=click&contentCollection=Personal%20Tech&module= RelatedCoverage&region=Marginalia&pgtype=article.   
Gillespie, Tarleton. ‘The Stories That Tools Tell’, in John Caldwell and Anna Everett (eds), New Media: Theses on Convergence Media and Digital Reproduction, New York: Routledge, 2003, pp. 107-123.  
\_\_\_\_\_. ‘The Politics of ‘Platforms’, *New Media & Society,* 12.3 (2010): 347-364.  
Gitelman, Lisa (ed.), *Raw Data is An Oxymoron*, Cambridge MA: MIT Press, 2013.  
Goel, Vindu. ‘As Data Overflows Online, Researchers Grapple with Ethics’, *New York Times*, 12 August 2014, http://www.nytimes.com/2014/08/13/technology/the-boon- of-online-data-puts-social-science-in-a-quandary.html?\_r=0.   
Government of Canada. *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*, 2014, http://www.pre.ethics.gc.ca/pdf/eng/tcps2- 2014/TCPS\_2\_FINAL\_Web.pdf.   
\_\_\_\_\_. ‘Tri-Agency Statement of Principles on Digital Data Management’, 21 December 2016, http://www.science.gc.ca/eic/site/063.nsf/eng/h\_83F7624E.html.   
\_\_\_\_\_. ‘DRAFT: Tri-agency Research Data Management Policy for Consultation’, 25 May 2018, http://www.science.gc.ca/eic/site/063.nsf/eng/h\_97610.html.   
Halpern, Orit. *Beautiful Data: A History of Vision and Reason Since 1945*, Durham: Duke University Press, 2015.  
Haggerty, Kevin D. ‘Ethics Creep: Governing Social Science Research in the Name of Ethics’, *Qualitative Sociology* 27.4 (2004): 391-414.  
Hauge Michelle V., Mark D. Stevenson, Kim D. Rossmo and Steven Le Comber. ‘Tagging Banksy: Using Geographic Profiling to Investigate a Modern Art Mystery’, *Journal of Spatial Science,* 61.1 (2016): 185–190.  
Hearn, Alex. ‘Google will Stop Scanning Content of Personal Emails’, *The Guardian*, 26 June 2017, https://www.theguardian.com/technology/2017/jun/26/google-will-stop- scanning-content-of-personal-emails.

\_\_\_\_\_. ‘How firms you have never interacted with can target your Facebook’, *The Guardian*. 21 April 2018, https://www.theguardian.com/technology/2018/apr/21/how- firms-you-have-never-interacted-with-can-target-your-facebook

Hill, Kashmir. ‘Facebook Added ‘Research’ to User Agreement 4 Months After Emotion Manipulation Study’, *Forbes*, 30 June 2014, www.forbes.com/sites/kashmirhill/2014/06/30/facebook-only-got-permission-to-do- research-on-users-after-emotion-manipulation-study.

Iliadis, Andrew, and Federica Russo. ‘Critical Data Studies: An Introduction’, *Big Data & Society*, December 2016, https://doi.org/10.1177/2053951716674238.

Kitchin, Rob. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*, London: SAGE, 2014.

\_\_\_\_\_. ‘Big Data, New Epistemologies and Paradigm Shifts’, *Big Data & Society* 1.1 (2014): 1-12.

\_\_\_\_\_, and Tracey Lauriault. ‘Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work’, *The Programmable City Working Paper 2* (2014), http://papers.ssrn.com/sol3/papers.cfm?abstract\_id=2474112

Kramer, Adam, Jamie Guillory and Jeffrey Hancock. ‘Experimental Evidence of Massive- Scale Emotional Contagion Through Social Networks’, *Proceedings National Academy of Science,* 111.24 (2014): 8788–8790.

Leathern, Rob. ‘A New Level of Transparency for Ads and Pages’, *Facebook Newsroom*, 28 June 2018, https://newsroom.fb.com/news/2018/06/transparency-for-ads-and-pages/.

Lee, Sandra Soo-Jin. ‘Studying ‘Friends’: The Ethics of Using Social Media as Research Platforms’, *The American Journal of Bioethics* 17.3 (2017): 1-2.

Luka, Mary Elizabeth, and Mélanie Millette. ‘(Re)framing Big Data: Activating Situated Knowledges and a Feminist Ethics of Care in Social Media Research’, *Social Media + Society*, April 2018, http://journals.sagepub.com/doi/full/10.1177/2056305118768297   
Markham, Annette. ‘Ethic as Method, Method as Ethic’, Journal of Information Ethics 15.2 (2006): 37–54.  
\_\_\_\_\_, and Elizabeth Buchanan. *Ethical Decision-Making and Internet Research 2.0:* *Recommendations* *From* t*he* *AoIR* *Ethics* *Working* *Committee*, 2012, www.aoir.org/reports/ethics2.pdf  
\_\_\_\_\_. and Elizabeth Buchanan. ‘Ethical Considerations in Digital Research contexts’, in James D. Wright (ed.), *Encyclopedia for Social & Behavioral Sciences,* Waltham MA: Elsevier, 2015, pp. 606-613.  
\_\_\_\_\_. ‘Afterword: Ethics as Impact-Moving from Error-Avoidance and Concept-Driven Models to a Future-Oriented Approach’, *Social Media + Society* (July-September 2018): 1-11.  
\_\_\_\_\_, Katrin Tiidenberg, and Andrew Herman. ‘Ethics as Methods: Doing ethics in the Era of Big Data Research-Introduction’, *Social Media + Society* (July-September 2018): 1-9, http://journals.sagepub.com/doi/abs/10.1177/2056305118784502   
Mayer-Schönberger, Viktor, and Kenneth Cukier. *Big Data: A Revolution that will Transform How we Live, Work, and Think*, Boston MA: Houghton Mifflin Harcourt, 2013.  
Metcalf, Jacob, and Kate Crawford. ‘Where are Human Subjects in Big Data Research? The Emerging Ethics Divide’, *Big Data & Society*, June 2016, http://journals.sagepub.com/doi/full/10.1177/2053951716650211.   
\_\_\_\_\_, Emily Keller, and danah boyd. *Perspectives on Big Data, Ethics, and Society*. Council for Big Data, Ethics, and Society, 23 May 2016, https://bdes.datasociety.net/council-output/perspectives-on-big-data-ethics-and- society/.   
Meyer, Robinson. ‘Everything we Know About Facebook’s Secret Mood Manipulation Experiment’, *The Atlantic*, 28 June 2014, https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about- facebooks-secret-mood-manipulation-experiment/373648/.   
Michael, Mike and Deborah Lupton. ‘Toward a Manifesto for the 'Public Understanding of Big Data’, Public Understanding of Science 25.1 (2015): 104–116.  
Napoli, Philip, and Robyn Caplan. ‘Why Media Companies Insist They're Not Media Companies, Why They’re Wrong, and Why it Matters’, ***First Monday*** 22.5 (2017).  
Neff, Gina, Anissa Tanweer, Brittany Fiore-Gartland and Laura Osburn. ‘Critique and Contribute: A Practice-Based Framework for Improving Critical Data Studies and Data Science’, *Big Data* 5.2 (2017): 85–97.  
Nissenbaum, Helen. *Privacy in Context: Technology, Policy and the Integrity of Social Life*, Palo Alto: Stanford University Press, 2009.  
\_\_\_\_\_. ‘A Contextual Approach to Privacy Online’, *Daedalus,* 140*.*4 (2011): 32–48.  
Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press, 2018.  
O’Neil, Cathy. ‘On Being a Data Skeptic’, O’Reilly Media, 2013, http://www.oreilly.com/data/free/files/being-a-data-skeptic.pdf.   
Penn State. ‘IRB Guideline X - Guidelines for Computer and Internet-Based Research Involving Human Participants’, Office of the Vice President for Research, 2018, https://www.research.psu.edu/irb/policies/guideline10.  
Plantin, Jean-Christophe, Carl Lagoze and Paul N. Edwards. ‘Re-Integrating Scholarly Infrastructure: The Ambiguous Role of Data Sharing Platforms’, *Big Data & Society* (January 2018), https://doi.org/10.1177/2053951718756683.   
Posner, Miriam, and Lauren F. Klein. ‘Editor’s Introduction: Data as Media’, *Feminist Media Histories* 3.3 (2017): 1-8.  
Reigeluth, Tyler Butler. ‘Why Data is Not Enough: Digital Traces as Control of Self and Self-Control’, *Surveillance & Society,* 12.2 (2014),

Rhodan, Maya. ‘“Please send help.” Hurricane Harvey victims turn to Twitter and Facebook’, *Time*, 30 August 2017, http://time.com/4921961/hurricane-harvey-twitter-facebook- social-media/   
Rieder, Gernot and Judith Simon. ‘Datatrust: Or, the Political Quest for Numerical Evidence and the Epistemologies of Big Data’, *Big Data & Society,* 3.1 (June 2016), http://journals.sagepub.com/doi/abs/10.1177/2053951716649398   
Rooke, Barry. ‘Four Pillars of Internet Research Ethics with Web 2.0’, *Journal of Academic Ethics,* 111.4 (2013): 265-268.   
Rusbridger, Alan and Ewen MacAskill. ‘Edward Snowden Interview – The Edited Transcript. *The Guardian*, 18 July 2014, https://www.theguardian.com/world/2014/jul/18/-sp- edward-snowden-nsa-whistleblower-interview-transcript.   
Sabati, Sheeva. ‘Upholding “Colonial Knowing” Through the IRB: Reframing Institutional Research Ethics’, *Qualitative Inquiry* (August 2018), DOI: https://doi.org/10.1177/1077800418787214.

Schroepfer, Mike. ‘An Update on Our Plans to Restrict Data Access on Facebook’, *Facebook Newsroom*, 4 April 2018, https://newsroom.fb.com/news/2018/04/restricting-data- access/   
Seetharaman, Deepa, and Georgia Wells. ‘Hurricane Harvey Victims Turn to Social Media for Assistance’, *The Wall Street Journal*, 29 August 2017, https://www.wsj.com/articles/hurricane-harvey-victims-turn-to-social-media-for- assistance-1503999001.  
Shilton, Katie. ‘Emerging Ethics Norms in Social Media Research’, *Big Data Ethics*, 2016 https://bigdata.fpf.org/papers/emerging-ethics-norms-in-social-media-research/.   
Shotwell, Alexis. *Against Purity: Living Ethically in Compromised Times*, Minnesota: University of Minnesota Press, 2016.  
Simmerling, Mary, Brian Schwegler, Joan E, Sieber and James Lindgren. ‘Introducing a New Paradigm for Ethical Research in the Social, Behavioral, and Biomedical Sciences: Part I’, Northwestern University Law Review 101*.*2 (2007): 837-858.  
Singer, Natasha. ‘What You Don’t Know About How Facebook Uses Your Data’, *The New York Times*, 11 April 2018, https://www.nytimes.com/2018/04/11/technology/facebook-privacy-hearings.html*.*    
Sonderby, Chris. ‘Reinforcing Our Commitment to Transparency’, *Facebook Newsroom*, 15 May 2018, https://newsroom.fb.com/news/2018/05/transparency-report-h2-2017/   
Stark, Laura. *Behind Closed Doors: IRBs and the Making of Ethical Research*, Chicago: University of Chicago Press, 2011.  
Sula, Chris Allen. ‘Research Ethics in an Age of Big Data’, *Association for Information Science and Technology*, 6 January 2016, https://www.asist.org/publications/bulletin/decemberjanuary-2016/research-ethics-in- an-age-of-big-data/.   
Taylor, Joanna, and Claudia Pagliari. ‘Mining Social Media Data: How are Research Sponsors and Researchers Addressing the Ethical Challenges?’, *Research Ethics,* 14.2 (October 2017): 1-39.  
Tene, Omer, and Jules Polonetsky. ‘Beyond IRBs: Ethical Guidelines for Data Research’, *Washington and Lee Law Review Online*, 72.3 (2016): 457-471, https://scholarlycommons.law.wlu.edu/wlulr-online/vol72/iss3/7/.  
Tilenberg, Katrin. ‘Ethics in Digital Research’,in Uwe Flick (ed.), *The SAGE Handbook of Qualitative Data Collection*. Thousand Oaks: SAGE, 2018, pp. 466-481.  
Townsend, Leanne, and Claire Wallace. ‘Social Media Research: A Guide to Ethics’, University of Aberdeen, 2015, https://www.gla.ac.uk/media/media\_487729\_en.pdf.  
University of Sheffield. The Ethics of Internet-based and Social Media Research, 2016, https://www.sheffield.ac.uk/polopoly\_fs/1.644904!/file/Report\_Ethics\_of\_Social\_Me dia\_Research\_Jul16.pdf.   
Van Dijck, Jose. ‘Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology’, *Surveillance & Society* 12.2 (2014), http://ojs.library.queensu.ca/index.php/surveillance-and- society/article/view/datafication.   
Verma, Inder M. ‘Editorial Expression of Concern and Correction’, *Proceedings National Academy of Science* 111.29 (2014): 10779, http://www.pnas.org/content/111/29/10779.1.   
Virtová, Tereza, Tereza Stöckelová and Helena Krásná. ‘On the Track of C/overt Research: Lessons from Taking Ethnographic Ethics to the Extreme’, *Qualitative Inquiry,* 24.7 (2018): 453-463.  
Walters, Chris. ‘Facebook’s New Terms of Service: ‘We Can Do Anything We Want with Your Content. Forever’, *Consumerist*, 15 February 2009, https://consumerist.com/2009/02/15/facebooks-new-terms-of-service-we-can-do- anything-we-want-with-your-content-forever   
Wang, Yilun, and Michal Kosinski. ‘Deep Neural Networks are More Accurate than Humans at Detecting Sexual Orientation from Facial Images’, *Journal of Personality and Social Psychology* 114.2 (2018): 246-257.  
Weaver, Matthew. ‘Facebook Scandal: I am Being Used as Scapegoat – Academic Who Mined Data’, *The Guardian*, 21 March 2018, https://www.theguardian.com/uk- news/2018/mar/21/facebook-row-i-am-being-used-as-scapegoat-says-academic- aleksandr-kogan-cambridge-analytica.  
Weindling, Paul. ‘The Origins of Informed Consent: The International Scientific Commission on Medical War Crimes and the Nuremberg Code’, *Bulletin of the History of Medicine* 75.1 (2001): 37-71.  
Williams, Matthew L, Pete Burnap and Luke Sloan. ‘Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users’ Views, Online Context and Algorithmic Estimation’, *Sociology* 51.6 (2017): 1149-1168.   
Zimmer, Michael. ‘Research Ethics in the Big Data Era: Addressing Conceptual Gaps for Researchers and IRBs’, School of Information Studies University of Wisconsin, Milwaukee, 2015,https://bigdata.fpf.org/papers/research-ethics-in-the-big-data-era- addressing-conceptual-gaps-for-researchers-and-irbs/.   
Zwitter, Andrej. ‘Big Data Ethics’, *Big Data & Society* (July 2014), <https://journals.sagepub.com/doi/abs/10.1177/2053951714559253>.

1. Australian National Data Service, *Data Sharing Considerations for Human Research Ethics Committees*, June 2018, <http://www.ands.org.au/guides/data-sharing-considerations-for-hrecs>; Ellen Broad, ‘Australia, we need to talk about data ethics’, *The Ethics Centre*, 25 January 2017, http://www.ethics.org.au/on-ethics/blog/january-2017/australia-data-ethics; dana boyd, ‘Where do we find ethics?’ *Points*, 5 April 2016, https:// points.datasociety.net/where-do-we-find-ethics-d0b9e8a7f4e6; Karin Clark, Matt Duckham, Marilys Guillemin, Assunta Hunter, Jodie McVernon, Christine O’Keefe, Cathy Pitkin, Steven Prawer, Richard Sinnott, Deborah Warr and Jenny Waycott, *Guidelines for the Ethical Use of Digital Data in Human Research*, The University of Melbourne, 2015, https://www.carltonconnect.com.au/read-ethical-use-of-digital-data-in-human-research/; Huw Davies et al, ‘Ethics Guidelines and Collated Resources for Digital Research’, British Sociological Association, 2017, https://www.britsoc.co.uk/media/24309/bsa\_statement\_of\_ethical\_practice\_annexe.pdf; Annette N Markham and Elizabeth Buchanan, ‘Ethical Considerations in Digital Research contexts’, in James D. Wright (ed.), *Encyclopedia for Social & Behavioral Sciences,* Waltham MA: Elsevier, 2015, pp. 606-613; Joanna and Claudia Pagliari, ‘Mining Social Media Data: How are Research Sponsors and Researchers Addressing the Ethical Challenges?’, *Research Ethics* 14*.*2 (October 2017): 1-39; Katie Shilton, ‘Emerging Ethics Norms in Social Media Research’, *Big Data Ethics*, 2016, https://bigdata.fpf.org/papers/emerging-ethics-norms-in-social-media-research/; Chris Allen Sula, ‘Research Ethics in an Age of Big Data’, *Association for Information Science and Technology*, 6 January 2016, https://www.asist.org/publications/bulletin/decemberjanuary-2016/research-ethics-in-an-age-of-big-data/; Katrin Tilenberg, ‘Ethics in Digital Research’,in Uwe Flick (ed.), *The SAGE Handbook of Qualitative Data Collection*, Thousand Oaks: SAGE, 2018, pp. 466-481; Leanne Townsend and Claire Wallace, ‘Social Media Research: A Guide to Ethics’, University of Aberdeen, 2015, https://www.gla.ac.uk/media/media\_487729\_en.pdf; University of Sheffield, *The Ethics of Internet-based and Social Media Research,* 2016, https://www.sheffield.ac.uk/polopoly\_fs/1.644904!/file/Report\_Ethics\_of\_Social\_Media\_Research\_Jul16.pdf. [↑](#footnote-ref-1)
2. dana boyd and Kate Crawford, ‘Critical Questions for Big Data: Provocations for a Cultural, Technological and Scholarly Phenomenon’, *Information, Communication & Society* 15.5 (2012): 662–679; Kate Crawford, Kate Miltner and Mary L. Gray, ‘Critiquing Big Data: Politics, Ethics, Epistemology’, *International Journal of Communication* 8(2014): 1663–1672; Viginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York: St. Martin’s Press, 2018; Andrew Guthrie Ferguson, *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*, New York: NYU Press, 2017; Andrew Iliadis and Federica Russo, ‘Critical Data Studies: An Introduction’, *Big Data & Society*, December 2016, https://doi.org/10.1177/2053951716674238; Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press, 2018; Miriam Posner and Lauren F. Klein, ‘Editor’s Introduction: Data as Media’, *Feminist Media Histories,* 3.3 (2017): 1-8; Jose van Dijck, ‘Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology’, *Surveillance & Society* 12.2 (2014). [↑](#footnote-ref-2)
3. Maya Rhodan, ‘‘Please send help.’ Hurricane Harvey victims turn to Twitter and Facebook’, *Time*, 30 August 2017, http://time.com/4921961/hurricane-harvey-twitter-facebook-social-media/; Deepa Seetharaman and Georgia Wells, ‘Hurricane Harvey Victims Turn to Social Media for Assistance’, *The Wall Street Journal*, 29 August 2017, https://www.wsj.com/articles/hurricane-harvey-victims-turn-to-social-media-for-assistance-1503999001. [↑](#footnote-ref-3)
4. Randall E. Bryant, Randy H Katz and Edward D. Lazowska, ‘Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science, and Society’, Computing Research Consortium, 22 December 2008,

   https://cra.org/ccc/wp-content/uploads/sites/2/2015/05/Big\_Data.pdf; Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution that will Transform How we Live, Work, and Think*, Boston, MA: Houghton Mifflin Harcourt, 2013. [↑](#footnote-ref-4)
5. Alan Rusbridger and Ewen MacAskill, ‘Edward Snowden Interview – The Edited Transcript. *The Guardian*, 18 July 2014, https://www.theguardian.com/world/2014/jul/18/-sp-edward-snowden-nsa-whistleblower-interview-transcript. [↑](#footnote-ref-5)
6. Sam Biddle, ‘Stop Using Unroll.me, Right Now. It Sold Your Data to Uber’, *The Intercept*, 24 April 2017, https://theintercept.com/2017/04/24/stop-using-unroll-me-right-now-it-sold-your-data-to-uber/; Alex Hearn, ‘Google will Stop Scanning Content of Personal Emails’, *The Guardian*, 26 June 2017, https://www.theguardian.com/technology/2017/jun/26/google-will-stop-scanning-content-of-personal-emails. [↑](#footnote-ref-6)
7. Sauvik Das and Adam Kramer, ‘Self-censorship on Facebook’, *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, 2013, pp. 120-127, https://research.fb.com/publications/self-censorship-on-facebook/; Michelle V. Hauge, Mark D. Stevenson, Kim D. Rossmo and Steven Le Comber, ‘Tagging Banksy: Using Geographic Profiling to Investigate a Modern Art Mystery’, *Journal of Spatial Science* 61*.*1 (2016): 185–190; Adam Kramer, Jamie Guillory and Jeffrey Hancock, ‘Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks’, *Proceedings National Academy of Science* 111.24 (2014): 8788–8790; Inder M Verma, ‘Editorial Expression of Concern and Correction’, *Proceedings National Academy of Science* 111.29 (2014): 10779;

   Yilun Wang and Michal Kosinski, ‘Deep Neural Networks are More Accurate than Humans at Detecting Sexual Orientation from Facial Images’, *Journal of Personality and Social Psychology* 114.2 (2018): 246-257. [↑](#footnote-ref-7)
8. Richard Adams, ‘Cambridge University asks Facebook for Evidence About Role of Academic’, *The Guardian*, 20 March 2018, https://www.theguardian.com/uk-news/2018/mar/20/cambridge-university-asks-facebook-for-evidence-about-role-of-academic-alex-kogan; Carole Cadwalladr and Emma Graham-Harrison, ‘Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach’, *The Guardian*, 17 March 2018, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election; Matthew Weaver, ‘Facebook Scandal: I am Being Used as Scapegoat – Academic Who Mined Data’, *The Guardian*, 21 March 2018, https://www.theguardian.com/uk-news/2018/mar/21/facebook-row-i-am-being-used-as-scapegoat-says-academic-aleksandr-kogan-cambridge-analytica [↑](#footnote-ref-8)
9. Natasha Singer, ‘What You Don’t Know About How Facebook Uses Your Data’, *The New York Times*, 11 April 2018, https://www.nytimes.com/2018/04/11/technology/facebook-privacy-hearings.html. [↑](#footnote-ref-9)
10. Mike Schroepfer, ‘An Update on Our Plans to Restrict Data Access on Facebook’, *Facebook Newsroom*, 4 April 2018, https://newsroom.fb.com/news/2018/04/restricting-data-access/. [↑](#footnote-ref-10)
11. Tim Adams, ‘Facebook’s Week of Shame: The Cambridge Analytica Fallout’, *The Guardian*, 24 March 2018, https://www.theguardian.com/technology/2018/mar/24/facebook-week-of-shame-data-breach-observer-revelations-zuckerberg-silence; Cadwalladr and Graham-Harrison, ‘Revealed: 50 Million Facebook Profiles Harvested for Cam bridge Analytica in Major Data Breach’; Keith Collins and Larry Buchanan, ‘How Facebook Lets Brands and Politicians Target You’, *The New York Times*, 11 April 2018, https://www.nytimes.com/interactive/2018/04/11/technology/facebook-sells-ads-life-details.html;

    Sheera Frenkel, Matthew Rosenberg and Nicholas Confessore, ‘Facebook Data Collected by Quiz App Included Private Messages’, *The New York Times*, 10 April 2018, https://www.nytimes.com/2018/04/10/technology/facebook-cambridge-analytica-private-messages.html?action=click&contentCollection=Personal%20Tech&module=RelatedCoverage&region=Marginalia&pgtype=article; Matthew Weaver, ‘Facebook Scandal: I am Being Used as Scapegoat – Academic Who Mined Data’. [↑](#footnote-ref-11)
12. Mark Andrejevic and Kelly Gates, ‘Big Data Surveillance: Introduction’, *Surveillance & Society* 12.2 (2014): 185-96; Tom Boellstorff, ‘Making Big Data, in Theory’, First Monday 18.10 (2013); Finn Brunton and Helen Nissenbaum, ‘Vernacular Resistance to Data Collection and Analysis: A Political Theory of Obfuscation’, ***First Monday*** 16.5 (2011); Finn Brunton and Helen Nissenbaum, *Obfuscation: A User’s Guide for Privacy and Protest*, Cambridge MA: MIT Press, 2016; danah boyd and Kate Crawford, ‘Critical Questions for Big Data: Provocations for a Cultural, Technological and Scholarly Phenomenon’; Kate Crawford, Kate Miltner and Mary L. Gray, ‘Critiquing Big Data: Politics, Ethics, Epistemology’; Lisa Gitelman (ed.), *Raw Data is An Oxymoron*, Cambridge MA: MIT Press, 2013; Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*, London: SAGE, 2014; Rob Kitchin, ‘Big Data, New Epistemologies and Paradigm Shifts’, *Big Data & Society* 1.1 (2014): 1-12; Jacob Metcalf, Emily Keller and danah boyd, *Perspectives on Big Data, Ethics, and Society*, Council for Big Data, Ethics, and Society, 23 May 2016, https://bdes.datasociety.net/council-output/perspectives-on-big-data-ethics-and-society/; Mike Michael and Deborah Lupton, ‘Toward a Manifesto for the Public Understanding of Big Data’, Public Understanding of Science 25.1 (2015): 104–116; Philip Napoli and Robyn Caplan, ‘Why Media Companies Insist They're Not Media Companies, Why They're Wrong, and Why it Matters’, ***First Monday*** 22.5 (2017); Gina Neff, Anissa Tanweer, Brittany Fiore-Gartland and Laura Osburn, ‘Critique and Contribute: A Practice-Based Framework for Improving Critical Data Studies and Data Science’, *Big Data* 5.2 (2017): 85–97; Cathy O’Neil, ‘On Being a Data Skeptic’, O’Reilly Media, 2013, http://www.oreilly.com/data/free/files/being-a-data-skeptic.pdf; Jose van Dijck, ‘Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology’. [↑](#footnote-ref-12)
13. Lisa Gitelman (ed.), *Raw Data is An Oxymoron*; Craig Dalton and Jim Thatcher, ‘What Does a Critical Data Studies Look Like, and Why Do We Care? Seven Points for a Critical Approach to “Big Data”’, *Society & Space Open Site* (2014), http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/; Craig Dalton, Linnet Taylor and Jim Thatcher. ‘Critical Data Studies: A Dialog on Data and Space’, *Big Data & Society* 3.1 (2014): 1–9; Rob Kitchin and Tracey Lauriault, ‘Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work’, *The Programmable City Working Paper* 2 (2014), http://papers.ssrn.com/sol3/papers.cfm?abstract\_id=2474112; Andrew Iliadis and Federica Russo, ‘Critical Data Studies: An Introduction’; Gernot Rieder and Judith Simon, ‘Datatrust: Or, the Political Quest for Numerical Evidence and the Epistemologies of Big Data’, *Big Data & Society,* 3.1 (2016). [↑](#footnote-ref-13)
14. Orit Halpern, *Beautiful Data: A History of Vision and Reason Since 1945*, Durham: Duke University Press, 2015. [↑](#footnote-ref-14)
15. Tyler Butler Reigeluth, ‘Why Data is Not Enough: Digital Traces as Control of Self and Self-Control’, *Surveillance & Society* 12.2 (2014), 249. [↑](#footnote-ref-15)
16. Mark Andrejevic and Kelly Gates, ‘Big Data Surveillance: Introduction’; Tyler Butler Reigeluth, ‘Why Data is Not Enough: Digital Traces as Control of Self and Self-Control’; Jose van Dijck, ‘Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology’. [↑](#footnote-ref-16)
17. Jacob Metcalf and Kate Crawford, ‘Where are Human Subjects in Big Data Research? The Emerging Ethics Divide’, *Big Data & Society* (June 2016), <http://journals.sagepub.com/doi/full/10.1177/2053951716650211>. [↑](#footnote-ref-17)
18. Jacob Metcalf and Kate Crawford, ‘Where are Human Subjects in Big Data Research? The Emerging Ethics Divide’, p. 3. [↑](#footnote-ref-18)
19. Mary Simmerling, Brian Schwegler, Joan E, Sieber and James Lindgren. ‘Introducing a New Paradigm for Ethical Research in the Social, Behavioral, and Biomedical Sciences: Part I’, Northwestern University Law Review 101*.*2 (2007): 837-858; Laura Stark, *Behind Closed Doors: IRBs and the Making of Ethical Research*, Chicago: University of Chicago Press, 2011; Paul Weindling, ‘The Origins of Informed Consent: The International Scientific Commission on Medical War Crimes and the Nuremberg Code’, *Bulletin of the History of Medicine* 75.1 (2001): 37-71. [↑](#footnote-ref-19)
20. Government of Canada, *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*, 2014, http://www.pre.ethics.gc.ca/pdf/eng/tcps2-2014/TCPS\_2\_FINAL\_Web.pdf, p.6. [↑](#footnote-ref-20)
21. Kirsten Bell, ‘Resisting Commensurability: Against Informed Consent as an Anthropological Virtue’, *American Anthropologist* 116.3 (2014): 511-522; Robert Dingwall, ‘The Ethical Case Against Ethical Regulation in Humanities and Social Science Research’, *Twenty-First Century Society* 3.1 (2008): 1-12; Malcom M. Feeley, ‘Legality, social research, and the challenge of institutional review boards’, *Law & Society Review* 41 (2007): 757-776. [↑](#footnote-ref-21)
22. Annette Markham, ‘Ethic as Method, Method as Ethic’, Journal of Information Ethics 15.2 (2006): 37–54, 39. [↑](#footnote-ref-22)
23. Kirsten Bell, ‘Resisting Commensurability: Against Informed Consent as an Anthropological Virtue’; Kakali Bhattacharya, ‘Consenting to the Consent Form: What are the Fixed and Fluid Understandings Between the Researcher and the Researched?’, *Qualitative Inquiry* 13.8 (2007): 1095-1115; Robert Dingwall, ‘The Ethical Case Against Ethical Regulation in Humanities and Social Science Research’; Malcom M. Feeley, ‘Legality, social research, and the challenge of institutional review boards’; Kevin D. Haggerty, ‘Ethics Creep: Governing Social Science Research in the Name of Ethics’, *Qualitative Sociology* 27.4 (2004): 391-414. [↑](#footnote-ref-23)
24. Kirsten Bell, ‘Resisting Commensurability: Against Informed Consent as an Anthropological Virtue’; Kakali Bhattacharya, ‘Consenting to the Consent Form: What are the Fixed and Fluid Understandings Between the Researcher and the Researched?’, Robert Dingwall, ‘The Ethical Case Against Ethical Regulation in Humanities and Social Science Research’; Malcom M. Feeley, ‘Legality, social research, and the challenge of institutional review boards’; Kevin D. Haggerty, ‘Ethics Creep: Governing Social Science Research in the Name of Ethics’; Sheeva Sabati, ‘Upholding “Colonial Knowing” Through the IRB: Reframing Institutional Research Ethics’, *Qualitative Inquiry* (August 2018), DOI: https://doi.org/10.1177/1077800418787214. [↑](#footnote-ref-24)
25. Government of Canada. *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*. [↑](#footnote-ref-25)
26. Ibid, p 3. [↑](#footnote-ref-26)
27. Government of Canada. *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*, p 6. [↑](#footnote-ref-27)
28. Ibid, p. 7 [↑](#footnote-ref-28)
29. Ibid, p.22. [↑](#footnote-ref-29)
30. Ibid, p.4. [↑](#footnote-ref-30)
31. Catherine Flick, ‘Informed Consent and the Facebook Emotional Manipulation Study’, *Research Ethics* 12.1 (2016): 14-28; Jacob Metcalf and Kate Crawford. ‘Where are Human Subjects in Big Data Research? The Emerging Ethics Divide’. [↑](#footnote-ref-31)
32. Adam Kramer, Jamie Guillory and Jeffrey Hancock, ‘Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks’. [↑](#footnote-ref-32)
33. Vindu Goel, “As Data Overflows Online, Researchers Grapple with Ethics”, *New York Times*, 12 August 2014, http://www.nytimes.com/2014/08/13/technology/the-boon-of-online-data-puts-social-science-in-a-quandary.html?\_r=0. [↑](#footnote-ref-33)
34. Robert Booth, ‘Facebook Reveals News Feed Experiment to Control Emotions’, *The Guardian*, 30 June 2014, https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds; Erin Egan, ‘Enhancing Transparency in Our Data Use Policy’, *Facebook Newsroom*, 11 May 2012, https://newsroom.fb.com/news/2012/05/enhancing-transparency-in-our-data-use-policy/; Robinson Meyer, ‘Everything we Know About Facebook’s Secret Mood Manipulation Experiment’, *The Atlantic*, 28 June 2014, https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/. [↑](#footnote-ref-34)
35. Catherine Flick, ‘Informed Consent and the Facebook Emotional Manipulation Study’. [↑](#footnote-ref-35)
36. Jacob Metcalf and Kate Crawford, ‘Where are Human Subjects in Big Data Research? The Emerging Ethics Divide’. [↑](#footnote-ref-36)
37. Sheeva Sabati, ‘Upholding ‘Colonial Knowing’ Through the IRB: Reframing Institutional Research Ethics’, *Qualitative Inquiry* (August 2018), DOI: https://doi.org/10.1177/1077800418787214, p.6. [↑](#footnote-ref-37)
38. Matthew L. Williams, Pete Burnap and Luke Sloan, ‘Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users’ View, Online Context and Algorithmic Estimation’, *Sociology* 51.6 (2017): 1149-1168. [↑](#footnote-ref-38)
39. Sandra Soo-Jin Lee, ‘Studying “Friends”: The Ethics of Using Social Media as Research Platforms’, *The American Journal of Bioethics* 17.3 (2017): 1-2. [↑](#footnote-ref-39)
40. Jacob Metcalf, Emily Keller, and danah boyd, *Perspectives on Big Data, Ethics, and Society*. [↑](#footnote-ref-40)
41. Leanne Townsend and Claire Wallace, ‘Social Media Research: A Guide to Ethics’. [↑](#footnote-ref-41)
42. Government of Canada, ‘Tri-Agency Statement of Principles on Digital Data Management’, 21 December 2016, http://www.science.gc.ca/eic/site/063.nsf/eng/h\_83F7624E.html. [↑](#footnote-ref-42)
43. Ibid. [↑](#footnote-ref-43)
44. Ibid. [↑](#footnote-ref-44)
45. Government of Canada, ‘DRAFT: Tri-agency Research Data Management Policy for Consultation’, 25 May 2018, http://www.science.gc.ca/eic/site/063.nsf/eng/h\_97610.html. [↑](#footnote-ref-45)
46. Matthew L. Williams, Pete Burnap and Luke Sloan, ‘Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users’ View, Online Context and Algorithmic Estimation’, p. 1150. [↑](#footnote-ref-46)
47. Government of Canada, *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*, p. 16. [↑](#footnote-ref-47)
48. Brian X. Chen, ‘I Downloaded the Information that Facebook has on Me. Yikes’, *The New York Times*, 11 April 2018, https://www.nytimes.com/2018/04/11/technology/personaltech/i-downloaded-the-information-that-facebook-has-on-me-yikes.html. [↑](#footnote-ref-48)
49. Ibid. [↑](#footnote-ref-49)
50. Huw Davies et al, ‘Ethics Guidelines and Collated Resources for Digital Research’; Charles Ess, *Digital Media Ethics*, Cambridge: Polity Press, 2014; Annette Markham, ‘Ethic as Method, Method as Ethic’; Annette Markham and Elizabeth Buchanan, *Ethical Decision-Making and Internet Research 2.0:* *Recommendations* *From the AoIR* *Ethics* *Working* *Committee*, 2012, www.aoir.org/reports/ethics2.pdf; Omer Tene and Jules Polonetsky, ‘Beyond IRBs: Ethical Guidelines for Data Research’, *Washington and Lee Law Review Online*, 72.3 (2016): 457-471, https://scholarlycommons.law.wlu.edu/wlulr-online/vol72/iss3/7/; Michael Zimmer, ‘Research Ethics in the Big Data Era: Addressing Conceptual Gaps for Researchers and IRBs’, School of Information Studies University of Wisconsin, Milwaukee, 2015, https://bigdata.fpf.org/papers/research-ethics-in-the-big-data-era-addressing-conceptual-gaps-for-researchers-and-irbs/; Andrej Zwitter, ‘Big Data Ethics’, *Big Data & Society* (July 2014),

    https://journals.sagepub.com/doi/abs/10.1177/2053951714559253. [↑](#footnote-ref-50)
51. Axel Bruns, ‘Faster than the Speed of Print: Reconciling ‘Big Data’ Social Media Analysis and Academic Scholarship’, *First Monday* 18.10 (2014); Katie Shilton, ‘Emerging Ethics Norms in Social Media Research’; Joanna Taylor and Claudia Pagliari, ‘Mining Social Media Data: How are Research Sponsors and Researchers Addressing the Ethical Challenges?’; Matthew L. Williams, Pete Burnap and Luke Sloan, ‘Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users’ View, Online Context and Algorithmic Estimation’. [↑](#footnote-ref-51)
52. Mary Elizabeth Luka and Mélanie Millette, ‘(Re)framing Big Data: Activating Situated Knowledges and a Feminist Ethics of Care in Social Media Research’, *Social Media + Society*, April 2018, http://journals.sagepub.com/doi/full/10.1177/2056305118768297; Jacob Metcalf, Emily Keller, and danah boyd, *Perspectives on Big Data, Ethics, and Society*, Council for Big Data, Ethics, and Society; Chris Allen Sula, ‘Research Ethics in an Age of Big Data’; Leanne Townsend and Claire Wallace, ‘Social Media Research: A Guide to Ethics’; University of Sheffield, The Ethics of Internet-based and Social Media Research. [↑](#footnote-ref-52)
53. danah boyd and Kate Crawford, ‘Critical Questions for Big Data: Provocations for a Cultural, Technological and Scholarly Phenomenon’. [↑](#footnote-ref-53)
54. Tyler Butler Reigeluth, ‘Why Data is Not Enough: Digital Traces as Control of Self and Self-Control’; Jacob Metcalf and Kate Crawford, ‘Where are Human Subjects in Big Data Research? The Emerging Ethics Divide’. [↑](#footnote-ref-54)
55. Michael Zimmer, ‘Research Ethics in the Big Data Era: Addressing Conceptual Gaps for Researchers and IRBs’, p. 6. [↑](#footnote-ref-55)
56. Clare Birchall, ‘Shareveillance: Subjectivity between Open and Closed Data’, *Big Data & Society* 3.2 (2016), https://journals.sagepub.com/doi/abs/10.1177/2053951716663965, p.5. [↑](#footnote-ref-56)
57. Ibid, p. 3. [↑](#footnote-ref-57)
58. An API is an interface that facilitates controlled access to the functionality and data contained by a software service or program, or in this case the Twitter platform. See Taina Bucher, ‘Objects of Intense Feeling: The Case of the Twitter API’, *Computational Culture* 3 (November, 2013), http://computationalculture.net/objects-of-intense-feeling-the-case-of-the-twitter-api/. [↑](#footnote-ref-58)
59. Tarleton Gillespie, ‘The Politics of “Platforms’”, *New Media & Society* 12.3 (2010): 347-364; Jean-Christophe Plantin, Carl Lagoze and Paul N. Edwards, ‘Re-Integrating Scholarly Infrastructure: The Ambiguous Role of Data Sharing Platforms’, *Big Data & Society* (January 2018), https://doi.org/10.1177/2053951718756683. [↑](#footnote-ref-59)
60. Jean-Christophe Plantin, Carl Lagoze and Paul N. Edwards. ‘Re-Integrating Scholarly Infrastructure: The Ambiguous Role of Data Sharing Platforms’. [↑](#footnote-ref-60)
61. Taina Bucher, ‘Objects of Intense Feeling: The Case of the Twitter API’. [↑](#footnote-ref-61)
62. Tarleton Gillespie, ‘The Politics of “Platforms”’; Tarleton Gillespie, ‘The Stories That Tools Tell’, in John Caldwell and Anna Everett (eds), New Media: Theses on Convergence Media and Digital Reproduction, New York: Routledge, 2003, pp. 107-123. [↑](#footnote-ref-62)
63. Tarleton Gillespie, ‘The Stories Digital Tools Tell’, p. 2. [↑](#footnote-ref-63)
64. Mark Andrejevic, ‘Big Data, Big Questions: The Big Data Divide’, *International Journal of Communication* 8 (2014): 1673-1689. [↑](#footnote-ref-64)
65. Kirsten Bell, ‘Resisting Commensurability: Against Informed Consent as an Anthropological Virtue’; Kakali Bhattacharya, ‘Consenting to the Consent Form: What are the Fixed and Fluid Understandings Between the Researcher and the Researched?’; Robert Dingwall, ‘The Ethical Case Against Ethical Regulation in Humanities and Social Science Research’; Malcom M. Feeley, ‘Legality, social research, and the challenge of institutional review boards’; Kevin D. Haggerty, ‘Ethics Creep: Governing Social Science Research in the Name of Ethics’; Sheeva Sabati, ‘Upholding ‘Colonial Knowing’ Through the IRB: Reframing Institutional Research Ethics’. [↑](#footnote-ref-65)
66. Annette Markham, ‘Afterword: Ethics as Impact-Moving from Error-Avoidance and Concept-Driven Models to a Future-Oriented Approach’, *Social Media + Society* (July-September 2018): 1-11. [↑](#footnote-ref-66)
67. Barry Rooke, ‘Four Pillars of Internet Research Ethics with Web 2.0’, *Journal of Academic Ethics*, 111.4 (2013): 265-268. [↑](#footnote-ref-67)
68. Barry Rooke, ‘Four Pillars of Internet Research Ethics with Web 2.0’; Katie Shilton, ‘Emerging Ethics Norms in Social Media Research’. [↑](#footnote-ref-68)
69. Jacob Metcalf and Kate Crawford, ‘Where are Human Subjects in Big Data Research? The Emerging Ethics Divide’. [↑](#footnote-ref-69)
70. Government of Canada, *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*, p. 14. [↑](#footnote-ref-70)
71. Penn State, ‘IRB Guideline X - Guidelines for Computer and Internet-Based Research Involving Human Participants’, Office of the Vice President for Research, 2018, https://www.research.psu.edu/irb/policies/guideline10. [↑](#footnote-ref-71)
72. Bloomberg Government, ‘Transcript of Zuckerberg’s Appearance Before House Committee’, *The Washington Post*, 11 April 2018, https://www.washingtonpost.com/news/the-switch/wp/2018/04/11/transcript-of-zuckerbergs-appearance-before-house-committee/?noredirect=on&utm\_term=.71d99a22271d. [↑](#footnote-ref-72)
73. Ibid. [↑](#footnote-ref-73)
74. Keith Collins and Larry Buchanan. ‘How Facebook Lets Brands and Politicians Target You’; Brian X. Chen, ‘I Downloaded the Information that Facebook has on Me. Yikes’; Sheera Frenkel, Matthew Rosenberg and Nicholas Confessore, ‘Facebook Data Collected by Quiz App Included Private Messages’; Natasha Singer, ‘What You Don’t Know About How Facebook Uses Your Data’. [↑](#footnote-ref-74)
75. Kashmir Hill, ‘Facebook Added ‘Research’ to User Agreement 4 Months After Emotion Manipulation Study’, *Forbes*, 30 June 2014, www.forbes.com/sites/kashmirhill/2014/06/30/facebook-only-got-permission-to-do-research-on-users-after-emotion-manipulation-study; Chris Walters, ‘Facebook’s New Terms of Service: ‘We Can Do Anything We Want with Your Content. Forever’, *Consumerist*, 15 February 2009, https://consumerist.com/2009/02/15/facebooks-new-terms-of-service-we-can-do-anything-we-want-with-your-content-forever. [↑](#footnote-ref-75)
76. Brian X. Chen, ‘I Downloaded the Information that Facebook has on Me. Yikes’; Keith Collins and Larry Buchanan, ‘How Facebook Lets Brands and Politicians Target You’; Alex Hearn, ‘How firms you have never interacted with can target your Facebook’, *The Guardian*. 21 April 2018, https://www.theguardian.com/technology/2018/apr/21/how-firms-you-have-never-interacted-with-can-target-your-facebook. [↑](#footnote-ref-76)
77. Helen Nissenbaum, *Privacy in Context: Technology, Policy and the Integrity of Social Life*, Palo Alto: Stanford University Press, 2009, p.127. [↑](#footnote-ref-77)
78. Helen Nissenbaum, ‘A Contextual Approach to Privacy Online’, *Daedalus* 140.4(2011): 32–48, https://www.amacad.org/publications/daedalus/11\_fall\_nissenbaum.pdf, p.38. [↑](#footnote-ref-78)
79. Helen Nissenbaum, ‘A Contextual Approach to Privacy Online’. [↑](#footnote-ref-79)
80. Helen Nissenbaum, *Privacy in Context: Technology, Policy and the Integrity of Social Life*; Helen Nissenbaum, ‘A Contextual Approach to Privacy Online’; Solon Barocas and Helen Nissenbaum, ‘Big Data’s End Run Around Procedural Privacy Protections’, *Communications of the ACM* 57.11 (2014): 31-33. [↑](#footnote-ref-80)
81. Orit Halpern, *Beautiful Data*. [↑](#footnote-ref-81)
82. Brian Barrett, ‘Facebook Owes You More Than This’, *Wired,* 19 March 2018, https://www.wired.com/story/facebook-privacy-transparency-cambridge-analytica/. [↑](#footnote-ref-82)
83. Chris Sonderby, ‘Reinforcing Our Commitment to Transparency’, *Facebook Newsroom*, 15 May 2018, https://newsroom.fb.com/news/2018/05/transparency-report-h2-2017/. [↑](#footnote-ref-83)
84. Rob Leathern, ‘A New Level of Transparency for Ads and Pages’, *Facebook Newsroom*, 28 June 2018, https://newsroom.fb.com/news/2018/06/transparency-for-ads-and-pages/. [↑](#footnote-ref-84)
85. Tereza Virtová, Tereza Stöckelová and Helena Krásná. ‘On the Track of C/overt Research: Lessons from Taking Ethnographic Ethics to the Extreme’, *Qualitative Inquiry* 24*.*7 (2018): 453-463. [↑](#footnote-ref-85)
86. Alexis Shotwell, *Against Purity: Living Ethically in Compromised Times*, Minnesota: University of Minnesota Press, 2016, p.107. [↑](#footnote-ref-86)
87. Alexis Shotwell, *Against Purity: Living Ethically in Compromised Times*, p.5 [↑](#footnote-ref-87)
88. Ibid. [↑](#footnote-ref-88)
89. Annette Markham, Katrin Tiidenberg and Andrew Herman, ‘Ethics as Methods: Doing ethics in the Era of Big Data Research-Introduction’, *Social Media + Society* 4.3 (2018): 1-9, 2. [↑](#footnote-ref-89)
90. Annette Markham, ‘Ethic as Method, Method as Ethic’, p.39; Annette Markham, ‘Afterword: Ethics as Impact-Moving from Error-Avoidance and Concept-Driven Models to a Future-Oriented Approach’. [↑](#footnote-ref-90)