(Re)framing Big Data: Activating Situated Knowledges and a Feminist Ethics of Care in Social Media Research

Mary Elizabeth Luka¹ and Mélanie Millette²

Abstract
In this article, we seek to problematize assumptions and trends in “big data” digital methods and research through an intersectional feminist lens. This is articulated through a commitment to understand how a feminist ethics of care and Donna Haraway’s ideas about “situated knowledge” could work methodologically for social media research. Taking up current debates within feminist materialism and digital data, including big, small, thick, and “lively” data, the argument addresses how a set of coherent feminist methods and a corollary epistemology is being rethought in the field today. We consider how the “queering” of Hannah Arendt’s concept of “action” could contribute to a critically optimistic and inclusive reflection on the role of ethical political commitments to the subjects/objects of study imbricated in big data. Finally, we use our recent research to pose a number of practical questions about practices of care in social media research, pointing toward future research directions.

Keywords
speculation, feminist materialism, Arendt, ethics of care, big data, situated knowledge

Introduction
To understand how ethics helps to problematize current fascinations with big data trends and their embedded biases, we first need to understand what is meant by data itself, both as a concept and as constructed, framed, pieces of information mobilized in our research. Data are commonly understood as the trace of an immediate relationship with a phenomenon. Merriam-Webster Dictionary defines data as “facts or information used to calculate, analyze, or plan something; information that is produced or stored by a computer” (“Data,” n.d.). Here lies the very heart of the problem: understanding data as a fact, or as zeros and ones, flattens their constructed, situated, and timely aspects. Consequently, the concept of data “remains categorically different from—and in a sense opposed to—the very idea of process.” (Markham, 2013, section 1). But even the most “immediate” data collection is the result of decisions made by researchers, computer analysts, and platform stakeholders (Bowker, 2006; Gitelman, 2013). When conducting research on Facebook, Twitter, Instagram, and other platforms, researchers mobilize software and hardware to capture social media data, scrape them, archive them, visualize them, and make sense of them. These different methodological steps and their results are shaped by decisions, biases, and values embodied by those who made them and those who disseminate or sell them (Crawford, 2016). Data are not neutral (Mulder, Ferguson, Groenewegen, Boersma, & Wolbers, 2016), and they cannot be an objective trace of a phenomenon, even in social media research where it is tempting to reduce a complex social experience to its digital traces.

Social media data can be sorted into different, sometimes complementary registers, depending on the research questions and goal, tools, and methodological strategies. “Big” data are commonly associated with quantitative strategies and are made distinct from the other registers by their massive volume, exhaustibility, indexability, and relationality (Kitchin & McArdle, 2016).¹ “Small” data, which are associated with qualitative methodology, are characterized by “a dataset composed of a relatively small collection of datapoints or cases, so that their analysis can be performed single-handedly via human coding and with little algorithmic assistance” (Latzko-Toth, Bonneau, & Millette, 2017, p. 202). “Thick” data come from ethnographic methods (e.g., Geertz, 1973) and are used

¹University of Alberta, Canada
²Université du Québec à Montréal, Canada

Corresponding Author:
Mary Elizabeth Luka, Campus Saint Jean, University of Alberta, 8406 Rue Marie-Anne Gaboury Northwest, Edmonton, Alberta, T6C 4G9, Canada.
Email: maryelizabethluka@ualberta.ca
to document complex and messy phenomenon like emotions or worldviews (Wang, 2016). Thick data build on small data sets by adding layers of data to each entry point of a small $n$, documenting the context, traces, and user’s experience (Latzko-Toth, Bonneau, & Millette, 2017). “Lively” data come from ordinary users’ daily online interactions and actions, as well as the automated analytical processes the production of those data engenders (Lupton, 2015).

In this article, we draw on our own research to unpack some of those implications and demonstrate the heterogeneity of ontologies of social media data. We aim to make explicit the differences among examinations of interlocked registers of big, thick, small, and lively data through our own case studies. Millette’s issue-based network analyses of thousands of tweets (“big” data) were followed by in-depth interviews (“thick” data), while Luka’s research involves the development of competing storytelling perspectives over a 4-year participatory media project (“thick” data) for the creation of a social media smart phone application (“small” data with the potential to generate “lively” and “big” data).

To develop our argument, we actively work toward operationalizing intersectional feminist perspectives in our roles as researcher citizens in a public sphere. We first discuss the concept of data to critically address their implications in research on social media. We then identify two of our own social media case studies before briefly reviewing an updated feminist “ethics of care” as we have mobilized it in that research. We draw from recent literature on feminist materialism to suggest how the concept of “speculation” may profitably be “turned” for use as ethical “practices of care” (methods), alongside aims to recognize intersectional situated knowledges in the social media context. Finally, we consider how such a methodology of speculation can be refined through Hannah Arendt’s notion of “action” in both the real and idealized world, even while acknowledging the challenges presented by her own epistemological and ontological framework. In this way, we aim to provide clear rationales and examples for conducting ethical social media research in the era of big data—and analyze our own actions as researchers.

**Critical Conceptualization of Data: Different Registers in Two Case Studies**

Geoffrey Bowker’s (2006) work on science, memory, and infrastructure offers a critical entry point into understanding how the collection and analysis of big, thick, small, and lively data are framed. More than 10 years ago, he showed how disciplinary memory is influenced by the way people deal with information. For example, researchers and assistants seek, document, and archive information differently. Each contribution shapes knowledge institutions disparately. To make this understood as an urgent ethical matter, he metaphorically noted that “Raw data is both an oxymoron and a bad idea; to the contrary, data should be cooked with care” (Bowker, 2006, p. 184). The phrase “cooking data” is commonly used in environments where interpretation is both needed and sometimes suspect, particularly including social media analysis, regardless of its source or the perspectives brought to such interpretations. Consequently, there is a cautionary tale to be understood by acknowledging the potential parallels between the idea of “cooking the data” and the idiom of “cooking the books,” a method of unethically altering financials and facts to a specific advantage.

Lisa Gitelman (2013) takes up the notion of understanding how we shape data ourselves by suggesting that each reassessment of data is the opposite of a direct engagement with reality. Of course, data can never fully represent reality, although data analyses provide pathways to help understand the world within which we live. Each time we analyze a data set, we impose assumptions that reshape it, even while the assumptions help respond to research questions. This has enormous implications for large, small, thick, and, especially, lively social media data sets. Deborah Lupton (2015) extends Gitelman’s argument by centering users in everyday data collection and consequent automated analyses. From her work on “data practices,” we can understand four characteristics of lively data:

1. Data are about life itself (health, transportation, interactions, emotions, memories, etc.);
2. Data are dynamic, always in the making, and constantly being re-organized by users themselves and different agencies;
3. Data play a central role in the knowledge economy, and global livelihood;
4. Data are very influential in everyday life at many levels (behaviors, beliefs, and decision-making).

We ourselves argue that these characteristics apply not only to social media data but also to “real life” data that we mobilize, for example, in our case studies. Lupton argues that online data are complex and develop simultaneously as self-tracking “data practices” that include self- or soft-surveillance, and through the ways in which these data are announced to users (push notifications, survey results, mobile adds, etc.). Lively data are paradoxical just as social media uses on major platforms have been shown to be paradoxical, especially for demographic minorities. Both contribute to one’s autonomy and resistance through self-expression and to alienation through self-surveillance, exploitation, and targeted advertisement (Proulx, Heaton, Choon, & Millette, 2011). Below, we make it clear that social media interactions are inextricably intertwined with other daily and historically shaped social relations, activities, and realities, likewise dynamic, influential, and reciprocal.

Both of the authors of this article have experienced the multi-faceted tensions of working in more than one register of data involving social media research. Luka is part of the research-creation group, Narratives in Space + Time Society
(NiS + TS), which uses social media, smart phone apps, augmented reality experiences, video, and public art walks (a combination of thick, small, and lively data) to respectfully articulate social tensions arising from conflicting racialized and classed histories in urban development. Most recently, Luka focused on the mobilization of thick data in the creation of an iOS app, “Drifts,” which makes space for the regeneration of data about social histories. The app structure and content enable and offer “thick description” (Geertz, 1973) by reframing the historical and social relations embedded in urban spaces related to the centenary of the 1917 Halifax Explosion, a wartime accident that injured or killed about 15% of the population of the Canadian port city (Luka & Lilley, in press). Millette conducted research with French-Canadian minority communities in English Canada, representing close to a million Francophones. Since Québec has a French-speaking majority, the Canadian province was left outside the research design to focus on political uses of social media by citizens in a minority language context. As a Québecker, Millette had to learn about the highly heterogeneous minority experience, and acknowledge the historical and political legacy that placed her in a privileged situation, to build trust with the actors. To conduct her research ethically, Millette opted for a critical approach inspired by ethnography to help develop a better sense of everyday life for the people engaged, online and offline (Hine, 2000). The approach resulted in an online immersion in blogs, Facebook groups, and Twitter feeds for more than 2 years. Millette (2015) performed “big” data mining on Twitter, but it was the two-and-a-half years of “thick” observation (on social media, as well as local newspapers and grassroots organization activities) and interviews that allowed her to understand and correctly code the thousands of collected tweets involved. In both the case studies, Luka and Millette exercised three core principles of critical research (based on George, 2014). First, each unpacked how knowledge is produced, making epistemologies, methodology, and biases explicit. Next, each revealed extant power relationships, in terms of both domination and agency. Finally, and simultaneously, each reported on what was regarded as unfair and unacceptable about the phenomenon under study and proposed solutions.

In both the case studies, the terrain was highly politically charged, broadly speaking, addressing desires to move beyond the colonial legacies of Canada. Each case involved sensitive historical elements and social relations. Each research design was developed to include the citizens (aka participants) concerned so that the research would give context to and reflect the complexity of the phenomena investigated. Each research team understood that the framing and analysis of mixed registers of data in a social media context are the result of processes where specific decisions are made and power relationships are continually negotiated, including through the selection of research objects, research teams, and the timing of data collection and analysis. Indeed, to resist the inappropriate “cooking” of data, each of these two research projects aimed to reframe dominant realities and perceptions, thereby counteracting established forms of social hierarchy, including when we worked on them or as we write them up for scientific papers. In the NiS + TS smart phone app, for example, the social co-collection, analysis, presentation, documentation, and subsequent representation of alternative and missing site-specific stories reveal the flattening of the data hierarchy that the NiS + TS public art walks enacted upon a previously relatively homogeneous and narrow data set: almost 100 years of colonialist storytelling about the 1917 Halifax Explosion. This opens the door to the collection of lively and big data in a social media context that is framed to recognize and analyze multiple perspectives, content, and approaches.

Annette Markham (2013) explains how such framing works by problematizing data from a Goffmanian perspective. Using the metaphors of frames and lenses, she illustrates how data and their representation have a symbolic power. This power is useful to situate ourselves in relation to a phenomenon, to focus on one aspect or tell-tale sign, and to ignore others. Yet, to forgive (or forget) what is left outside the frame is problematic. As with Bowker, Gitelman, and Mauthner, Markham argues that understanding data as “normal,” “natural,” and outside its own reality as a human-shaped tool for sense-making is a major concern in social sciences and science. Consequently, she calls for “more nuanced frameworks for what inquiry means as a process” (Markham, 2013, Section 5). Millette (2015) experienced such reframing many times while conducting her research about how Francophone minorities in Canada use social media to gain visibility and make political claims. One of the most significant examples occurred during her initial contacts with users. She used the phrase “the Francophone community” to describe their situation. Right away, participants responded by pointing out the homogeneity of such framing: there is no such thing as a (singular) Canadian Francophone community. The notion of “Canadian Francophones” turned out to be the most efficient expression for what inquiry means as a process (Markham, 2013, Section 5). Millette initially assumed she was observing and what the other participants were experiencing were different. By accepting the correction offered by the participants involved, that is, by researching with care, Millette’s research was strengthened and made more accurate.

**A Feminist Ethics of Care as a Framing Strategy for Ethical Practices in Research**

The sooner we understand social media data—big, small, thick, or lively—as a humanly constructed artifact shaped by power relationships and crafted according to certain values and standing points, the more we can realize how these processes enable and shape our decisions about our methodology. Both of our examples build on Annette Markham’s (2013) framing strategies and our own extension of Deborah
Lupton’s (2015) notion of “lively” data, over time and through the “real” (materialized) worlds of situated knowledge (Haraway, 1988). This allows us to draw on an updated ethics of care to illustrate how different kinds of data—implicated at different registers of engagement over time—can “turn” us in practical ways to critically rethink the ongoing intersectional networks of relations, values, and ethical commitments that undergird our research and those of others. In other words, building on Gitelman’s (2013) historical analysis, we are not simply suggesting that data take a different form and have a different value depending on methodological preferences and dominant epistemologies—but also, we would add, by daily practices and timings (Lupton, 2015), funding sources and ethical review board assumptions (Luka, Millette, & Wallace, 2017), and the ethical shareability of data (Mauthner, 2012).

Mauthner (2012) elucidates ethical tensions relating to “our moral responsibilities and commitments to respondents, [and] our moral ownership rights over the data we produce” (p. 158) to describe increasing tensions such as maintaining confidentiality of data versus the growing number of public funders or private marketers who ask researchers to make social media data sets reusable and shareable.

Such pressures support the importance of a commitment to developing caring frameworks of enquiry, and the practices that accompany them, as we have previously argued (Luka et al., 2017), particularly in the growing field of critical data studies, within which our argument is positioned. A contemporary “ethics of care” emerges from work initially undertaken in the 1980s onward, particularly that of second- and third-wave feminist psychology, sociology, and cultural studies. While grounded in the ethnographic work of Carol Gilligan (1982) and Norman K. Denzin (1997), for example, today’s ethics of care finds mature expression in works by Angela McRobbie (2015) and Miller, Birch, Mauthner, and Jessop (2012), by emphasizing the integration of feminist and intersectional values into considerations of data analyses, including big data. Key among these include identifying and respecting diversity, paying attention to how our research may affect those under study, and articulating and acknowledging our intent as researchers and participants, including whether and how we aim to generate potentially transformative engagements (Edwards & Mauthner, 2012). More specifically, Gillies and Alldred (2012) argue that our present-day ethics of care is built on three formative processes: “representing women” (p. 49) across colonial, developing, and indigenous points of view and experiences; “deconstructing and undermining ‘knowledge’ structures” (p. 55); and facilitating practices of care such as “initiating personal change through action research” (p. 51), as we demonstrate in this article. Intersectional feminist practices of care help us to identify biases and assumptions and to better understand what our commitments and objectives are, as well as those of our participants, partners, and others involved in the research process.

For example, Anna Lauren Hoffmann (2016, 2018) employs an ethics of care in critical data studies by demanding that data studies must be based on ethical concerns, and what she identifies in her own research as “data violence” conducted on subjects. Such violence can be physical or symbolic or both and is the result of data producing or being used to reproduce dominant norms. More specifically, she explains that data create and maintain distinctions based on colonial ideologies, in particular because they become part of our contemporary and daily practices, and help to legitimate discrimination. To illustrate how violence is practiced in this context, Hoffmann points to how we arbitrarily code the gender of human subjects, for example, in airport body scans. These machines assign binary gender identifications to blur the sexual organ zone of the body, as well as the breasts of “female”-identified travelers. This sociotechnical assignment is experienced as a violent act by trans and gender-fluid persons. Similarly, Rena Bivens (2015) analyzes the subterfuges practiced by Facebook in its assignment of gender binaries. While users have a range of options for the selection of gender identification, they are required to select preferred pronouns (he or she or they), which make their earlier gender identification moot, at least in terms of advertising targeting, and so on—one of the primary revenue-generating functions of many social media software.

However, Hoffmann (2016) suggests that even if data can be violent, they can also contribute to a more equitable society. How can we achieve such an apparently paradoxical state? How can data maintain complexity and remain situated? Similarly, Gibson-Graham (2015) asks how we can reimagine research as a set of ethical practices “rather than structural dynamics” (p. 57). Harking back to third-wave feminist understandings of multiple and fluid identities, these practices require mutual commitment and care in research, through an ongoing negotiation of power relations. Such an approach can ameliorate the tensions between demands for open data ongoing calls to protect privacy in data collections, including the ways in which explicit consent is neither sought nor secured for cyclical or repeated market or other research conducted via social media applications, also a potential act of violence for the participants involved.

In the case of the NiS+TS smart phone app, Luka and the rest of the research team decided to limit the type and specificity of analytics generated from the app. Although users of the app are welcome to provide their own user-generated content (comments, photos, and videos), they are able to set up their user name without that information being collected alongside the data aggregated about the volumes of users at specific points of interest on each of the walks (“drifts”). Maximizing the possibility of anonymity is regarded to be a priority by the entire research team, especially given the sensitivity of some of the stories told. In other words, how do we share, teach, and learn from best practices about data ownership, privacy rights, and responsibilities? For some users, and in our research, especially with inexperienced social media users, privacy is misunderstood or poorly valued. It is,
therefore, our responsibility as researchers to carefully anonymize the data collected, including avatars and profile images, and share the reasons why with participants. Such an approach is supported by international negotiations of data ownership conditions.

such as the empowerment of individuals, informational self-determination, non-discrimination, freedom, dignity, and autonomy need to be particularly taken into account in the context of personal data (management). Other values worth considering are the protection of the individual from physical and financial harm and the unhindered distribution of knowledge (Dell’Aglio et al., 2017, Section 1.4).

In contrast, sometimes identifying the user is crucial to their empowerment. For example, in her research with Canadian Francophones, Millette asked permission to identify an online group of citizens, TaGueule! (which means ShutUp!). This group represented a highly politicized and well-informed online public. Citing this collective was important as a gesture to recognize their efforts to bring attention to political claims about minority French culture and language in Canada (see http://www.taigueule.ca). Consequently, Millette (2014, 2015, 2016) cites their manifesto and many of their online posts as a meaningful and scientifically relevant source of knowledge in different academic works in coherence with what Donna Haraway (1988) calls “situated knowledges.”

**Situated Knowledge: Speculation as Ethical Method**

To more clearly unpack how social media data ontologies are preframed (and hold the potential to be reframed, in Markham’s terms), we suggest (alongside Asberg, Thiele, & van der Tuin, 2015) that feminist materialism can usefully reconceptualize speculation as method rather than ontology per se by way of Haraway’s (1988) notion of situated knowledges. Here, we refer to Haraway’s (1988) critique of the chimera of one unified knowledge (namely, the knowledge accepted by those in privileged position) and her call (still relevant today) for academic research methodologies to give us the “ability to partially translate knowledges among very different—and power-differentiated—communities” (p. 580). In this sense, “speculation” bears a relationship to its dictionary definition as “the contemplation or consideration of some subject: [e.g.] to engage in speculation on humanity’s ultimate destiny” (Dictionary.com, consulted 5 January 2018) or in its synonymous sense as forms of contemplation or reflection, or as conception, deliberation, or conjecture. Speculation as a method is not a set of rules but a gathering of possibilities. Speculation establishes the researcher’s stance as one among many, sensitive to their own biases, power, and identity positions, thereby compelling each to unpack or even give up preexisting conceptual frames and labels. In this context, we agree with Asberg et al. that speculation is both a way to think through these multiple (sometimes contradictory) ways to understand or address an issue and an optimistic gesture that can imagine alternative sets of social relations. Speculation helps shift the center of gravity in research. Researchers can thereby acknowledge their own role in shaping research by the data they choose and operationalize their own social commitments to equity and other social values by understanding the registers of data they participate in reproducing as well as study. Similarly, big data can be superseded, supplanted, or augmented by small, lively, or thick data (or vice-versa), in a shifting realm of changing conditions and overlapping ontologies of data, all within the span of any given project.

To ground the use of speculation methodologically, Asberg et al. explicitly make use of situated knowledges to ensure that the subject is not transcended by the object of study (data), instead recognizing the codependence of object and subject. As in the work of Hoffmann, (2016), Bivens (2015), and Roberts (2016), Asberg et al. recognize how knowledge is pre-integrated at every turn. They “... insist on the co-constitutive role of the embedded observer, the perspective and the rich agentiality (multi-subjectivity) of context itself...” (p. 151). Such an approach deeply informed the methods used by NiS + TS to develop its social media app and shift a seemingly immutable history about the Halifax Explosion, including ongoing documentation of lively and thick data, such as collaborative research walks, workshops with neighborhood participants, artist collaborators involved in interpreting and presenting stories, and the documentation and representation of that data in the smart phone app. The app continues to collect yet another layer of lively social media data through comments, videos, audio recordings, and photographs shared to the app. For social media in the age of big data, activating research with others rather than conducting it upon others is a material act of performative speculation which “envision[s] a different world and challenge[s] taken-for-granted knowledges by way of situating them in specific historical, sociocultural, material and bodily contexts” (Asberg et al., 2015, p. 153).

Ethically, such an approach reflects the practices of care in research that we argue are part of today’s research obligations from an intersectional feminist perspective. In the development of the NiS + TS app, it was important to provide user-generated content (UGC) capabilities and to ensure that each round of data collection would be continuously reshaped by the prioritizing of heretofore missing narratives and perspectives. Several sites (“points of interest”) in the walks (“drifts”) explicitly recentre Mi’kmaq and African-Canadian communities and histories, especially when visible evidence has long since been erased through urban development and discriminatory resettlement. Stories of erasure and recovery are told by Mi’kmaq and African-Canadians, not just as counterpoints to more traditional versions of the Halifax Explosion narratives but as new centers. Such research reflects what Asberg et al. name the penultimate stage of
speculative practice, in which researchers become more skilled at identifying and working with contextual circumstances to articulate appropriate situated knowledges. Such a proposition gives over the space required to activate a more equitable and less institutionalized world. Asberg et al. further insist that feminist speculative practices of care must unpack deeply embedded binary categories of data in many social sciences, arts, and humanities traditions, including studies of big data derived from social media. This suggests “the need to move away from pre-givenness . . . towards a fuller, more complex and surprising world [that cannot be understood through binaries such as] subject (Thought) and object (Being)” (Asberg et al., 2015, p. 161). It is not just that intersectional feminist practices of care contest what Haraway (1988) called “the god trick”—findings that assert that the world is the way that it is because it has always been that way, or a world divided into typologies based on absolutist binaries. It is also that we can move beyond binaries, including theoretically foundational notions such as the division of private versus public spheres. Indeed, some authors (including Benhabib, 1993; Jones, 2015) suggest that self-imposed limitations based on the private/public binary can be disrupted through materialist speculation methodologically. It is to Arendt’s notion of “action” that we now turn to test this idea.

Arendt’s Paradoxical Legacy: Beyond the Citizen/Researcher Binary

Revisiting the humanist contribution of Hannah Arendt’s (1961) notion of “action” is helpful to epistemologically ground an intersectional feminist ethics of care, responding to our researcher identities and connecting us more broadly as citizens and political actors in the world. Arendt’s definition of public (vs private) space is both critical and utopic in this context—and has been a puzzling object of discussion among feminists for a long time.6

More relevant to our argument in this context, Arendt (1961) defines action as a radically free activity. Her concept demands more than just “taking a decision”; for Arendt, a decision follows from reason, “a judgment from the intellect” and will (p. 152). In contrast, freedom is a speculative or aspirational principle, which inspires the highest human actions—such as art and politics—and can precede or even supercede so-called “rational” thinking. Aware of the challenging abstraction inherent in her conception of freedom, Arendt (1961) uses the virtuosity and ephemerality of artistic performance as a metaphor to explain its power. She suggests that the difference between action and potential activity is the performance of action, or the difference between a concert performance and the existence of a written concerto.

This understanding of Arendt’s action resonates strongly with speculation as method, especially by offering possibilities for a more inclusive society. For example, Kathleen B. Jones (2015) suggests that we rethink Arendt’s relationship with intersectional feminism by considering how Arendt’s action supports coalition-building to disrupt oppression and power relations, noticeable through identity categories, such as gender, race, class, and nation (p. 459). Jones suggests that by reframing Arendt’s notion of “action,” we can engage with how public and private spheres intersect with and enact intersectional feminism, “queer-ing” and “queering” it in the process, as a form of resistance. “Queering” in this context asks us to analyze unacknowledged costs involved in data production and analysis and make visible what is invisible or otherwise discounted. For example, the work of Sarah Roberts (2016) in identifying the traumatic social and emotional costs borne primarily by Asian, Filipino, and Indian laborers who are hired to review contested content in mostly Western social media forums such as Facebook, Twitter, and Instagram explicitly “queers” our understanding of social media as a platform for free expression and democratic practices. Such queered analysis speculatively reframes perspectives in research. Consequently, Roberts helps us to unpack the failings of neoliberalism, including specious claims of new forms of automation, and to better understand how and why ubiquitous uses of social media—combined with physical migrations and diasporic movements—impact us all. These types of questions are used in research justice methodologies that reveal and support social change, recentering previously marginalized groups and thereby generatively “queering” how researchers and writers work together to publish work (e.g., Zefferio & Hogan, 2015). Indeed, like Hoffmann (2016), Jones argues that the ability to understand such phenomena in the world can go through a kind of “reverse engineering” that mobilizes alternative approaches, political actions, and ethical practices (Jones, 2015, pp. 460-461)—or, as Arendt might put it, activating action.

From an intersectional feminist point of view, we must ask how our recognition of others (and others of ourselves) can be reconciled within the limits of dominant, heteronormative, colonial norms and frames, including those resisted by Arendt. How can a subject acquire legitimacy, without recourse to binaries or being co-opted by dominant standards or without using dominant narratives, institutions, and so on, which are often oppressive and inaccurate for those who embody “difference”? In social media research, this can be achieved not just by building content or surveys together with participants or by explaining the presence of participants in developing proposals for funding, and so on. In Millette’s research, for example, her methodological design included rounds of interviews at the end of specific research phases to verify whether the research team had interpreted the community’s online traces on social media correctly, clarifying and validating certain interpretations derived from Millette’s analysis of the relevant Twitter data.

Arendt (1958) elaborates the common world as the condition for a public political space, which needs the co-recognition
(or, in feminist materialist terms, the co-constitutive nature) of participating subjects to exist. In Arendt’s terms, a public, democratic engagement emerges when various subject-positions meet each other, and authentically share their realities. The importance of co-recognition in Arendt’s theory is more explicitly addressed in the work of Jones on the nature of how we relate to others. She argues that Arendt “learned how important it is that other people, who live in the same historical space as you do, recognize who you are, see you as the person you understand yourself to be, and accept you as you are” (Jones, 2015, p. 462). In Luka’s research, this is expressed through the participation of more than 80 collaborators in the reframing of historical tropes and social spaces as content for the NiS+TS app, as well as the structure within which that content is presented and shared with thousands of social media users. The process requires an ongoing negotiation of both time and historical understandings, sometimes achieved by literally (embodied-ly) walking through neighborhoods at appointed times, or visiting archives, or mounting exhibitions, or viewing and conducting interviews together. In Millette’s research on political uses of social media in French Canada, such embodiment was similarly achieved by traveling to meet users to interview them. Inspired by ethnographic practices (Geertz, 1998), Millette spent days in each city to hang out in cafés and other hubs identified as “Francophone-friendly.” Her goal was to conduct interviews, but this allowed her to experiment with what it meant to live in such environments and to develop a feeling for the cultural and linguistic differences. In the end, “hanging around” was key to the quality of the interviews. People she met told her, “we’ve seen you around for a while online, and now you are here. It seems like you really care.” Such comments illustrate the ways in which the generating of lively data takes place in person in combination with online activity. In the eyes of the Francophones, this everyday commitment was instrumental for legitimizing Millette’s commitment to conduct her research ethically. Similarly, Jones points out that “Arendt insisted ‘one can never and must never legitimize oneself. It is always others who do the legitimizing. And not just any others . . . but those who live in the same historical space’” (Arendt quoted in Jones, 2015, p. 462). Mutual recognition is central to the emergence of a common world, without which a public sphere is not possible, nor is political action. From this perspective, Arendt asks us to become radically intelligent to grasp the opportunity, the right moment” (Poizat, 2007, Para 32, interviewing Tassin; our translation). Thinking about methodology as “action” means that research is about process, and the deployment and analysis of data are also action—often with others—rather than outcome. In Arendt’s terms, conceptualizing freedom as a speculative principle for action would mean that aspiring to freedom can survive the “success” or “failure” of any singular activity related to realizing freedom. This is remarkably similar to the way in which feminist materialism employs speculation as method. In the case of the Drifts app, this was illustrated by what NiS+TS came to call “choreographing” rather than “scripting” either the public art walks or the app “drifts” that emerge from the walks. By generating content and frameworks for nine pliable, thematic walks together with our collaborators, NiS+TS created conditions for social media users to influence the walks, and social relations, in those neighborhoods, by adding UGC, by leading the walks themselves, or by building new walks. In Arendt’s (1958) words, since humans are capable of action, then the speculative or “the unexpected, can be expected” (p. 178).

Concluding Remarks: Ethical Practices of Care in Social Media Research

To conclude our discussion, we return to the practicalities required of a methodological enquiry. We have developed a preliminary set of questions that we have found useful in our own social media research endeavors, in combination with, for example, the professional standards articulated by the Association of Internet Researchers (AoIR) respecting quantitative and qualitative data sets (Markham & Buchanan, 2012). Since most social media environments are sensitive terrains for someone (if not everyone) involved, these questions emerge from our commitments to an ethics of care, through practices of care. They may help other researchers to benefit from a more responsive, ethical methodological design that recognizes participants’ needs and rights much more explicitly:

1. What right do we have to be on this terrain?
2. From which stance or positionality do we undertake our research and with which privileges and vulnerabilities?
3. Are we participants as well as researchers? In what ways?
4. How can this research be collaborative and useful for the person(s) whose experiences are being scrutinized? In what ways and to what degree are we comfortable inviting members of the communities involved to collaborate on research design?
5. What if they aren’t interested in explicit participation? How do we provide implicit decision-making opportunities through which participants can be

Our reframing of action, speculation, and ethics of care asks researchers to conduct any research as if it was critical action-research meant to contribute to a better, more equitable world. This could mean resisting the pressures of normatively acceptable research methods, to include social media users in research design. French philosopher Étienne Tassin explains that Arendt’s definition of freedom as the principle for action lies exactly in this mix of performative “courage and determination, audacity and circumspection, lucidity and intelligence to grasp the opportunity, the right moment” (Poizat, 2007, Para 32, interviewing Tassin; our translation). The importance of methodology as “action” means that research is about process, and the deployment and analysis of data are also action—often with others—rather than outcome. In Arendt’s terms, conceptualizing freedom as a speculative principle for action would mean that aspiring to freedom can survive the “success” or “failure” of any singular activity related to realizing freedom. This is remarkably similar to the way in which feminist materialism employs speculation as method. In the case of the Drifts app, this was illustrated by what NiS+TS came to call “choreographing” rather than “scripting” either the public art walks or the app “drifts” that emerge from the walks. By generating content and frameworks for nine pliable, thematic walks together with our collaborators, NiS+TS created conditions for social media users to influence the walks, and social relations, in those neighborhoods, by adding UGC, by leading the walks themselves, or by building new walks. In Arendt’s (1958) words, since humans are capable of action, then the speculative or “the unexpected, can be expected” (p. 178).
involved? If they accept the opportunity (and responsibility) to work with us on research design, how do we resolve any conflicts or disagreements? To what degree are we prepared to rethink our research plans and repurpose it to align with their agenda and needs?

6. How can we make our proposals to collect data visible to those whose work we wish to collect? danah boyd and Kate Crawford (2011) rightly insist that so-called “public” tweets, posts, and snaps are not research-ready; they were not originally meant to be research material—How can we help participants and other researchers to understand this and promote informed consent in that process?

7. How can we advance debate about issues of privacy and personal data ownership in a social media context? In a connected world where Google, Amazon, Facebook, and Apple (GAFA) form an oligopoly that largely dominates global data management in Western societies, how can we, as educators and theorists, take action? Are there workshops, alternative platforms or other processes that we can support and demonstrate through our research?

8. How do we disseminate the results in a respectful and useful way? If we are required to publish in journals for peers, how can we make this meaningful for the people whose involvement has contributed to these results? What are the other forms of reporting back, providing joint attribution or other analyses or descriptive expressions that we could commit to and involve participants in? Is a written format the best way, perhaps in a blog or newspaper? Would they prefer to be involved in an oral discussion (Online? In person?) A workshop? A collective art piece?

9. How will shared involvement improve our research? What if it doesn’t?

10. In what ways will our research contribute to a more equitable world, even if it is challenging or hard for us or others to imagine, or seems over-simplified or arrogant to others? In other words, how do we act speculatively to embody action?

In this article, our argument favors a complex understanding of big data and also of small, thick, or lively data—in a social media environment as human-shaped artifacts, which we suggest calls for a consciously feminist ethics of care, and a flexible set of practices of care. This is a fruitful way to address today’s ethical research challenges. Elsewhere, we (Luka et al., 2017) have shown that even a brief survey of recent feminist digital research reveals a focus on methods and processes that suggest the vibrancy of the disciplinary terrain through which the ethics of social media studies can both expand and become more ethically responsive. In this article, we have used some specific examples from our own recent social media research to illustrate how an ethics of care can be implemented in practice. Here, we have presented opportunities to revisit the activation of situated knowledge and ethical practices of care to interrogate the kind of social media research that is valorized in the era of “big data.” We argue that feminist materialist speculation as methodology, combined with the interpretive power of Arendt’s concept of action, can not only queer the research agenda of big data but may be one of the methodological answers to the complex challenges of conducting ethical and equitable research today, throughout different registers of data. We have offered insights based on our research to illustrate, methodologically, how to embody such a commitment. Of course, there are limits to the way these ethical practices of care can be applied. For example, what happens with gargantuan “big data” sets? What works in the examples we give might not work in other circumstances. Academia is under significant political and economic pressure to create jobs, engender social impacts, and balance books, as with many other sectors. And many researchers, especially those who are precariously employed, may not have the resources to apply what we propose here. We would argue, however, that many scholars occupy the kinds of privileged positions that allow for the enactment of ethical practices of care in any register of data research, and indeed, many do so. Even without unlimited—or even generous—means, researchers can still be cautious about the data gathered, the correlations found, the interpretations drawn, and our shared responsibility as participants in social relations, including social change.

We understand our role here to be those of citizens as well as researchers. From an intersectional feminist point of view, we have taken the opportunity to proffer the kind of research which could bring us closer to a robust speculative methodology that critically engages social media data and acknowledges—and involves—the people and social relations imbricated therein. From our perspective, a contextualized ethical approach to analyzing social media data—in part by taking up situated knowledges—are optimistic gestures to support those who fight to make the productively unexpected happen, including surfacing unseen or incompletely understood stories, conditions of living, and social relations in general. In Arendt’s (1958) words, such action must always be attempted because “the smallest act in the most limited circumstances bears the seed of the same boundlessness, because one deed, and sometimes one word, suffices to change every constellation” (p. 190).

Acknowledgements

Thank you to the reviewers who took the time to provide us with such constructive feedback; our article is greatly improved by your contribution. Thank you, also, to the guest editors whose work ensures that ours is placed into such a crucial dialogue at the right moment. And thank you to the participants in our research, without whom none of our work would have taken place.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article. The authors received no direct financial support for the research, authorship, and/or publication of this article. Each was supported by research fellowships to conduct elements of their own research projects discussed in the article, including the Social Sciences and Humanities Research Council Banting Postdoctoral Fellowship (Luka) and the Trudeau Foundation Doctoral Scholarship (Millette).

Notes
1. Kitchin and McArdle (2016) conducted an analysis of 26 published scientific studies described by their authors as “big data” studies. They used nine commonly established characteristics of big data (volume, velocity, variety, resolution, indexicality, relationality, extensionality, exhaustivity, and flexibility) to verify the consistency of the data sets collected and mobilized in the studies. They concluded that these characteristics vary a lot from one data set to the other, making it hard to describe what an ontology of big data would be. Ironically, the first “three Vs” of big data (volume, velocity, and variety) are among the less consistent characteristics. Their results suggest that velocity and exhaustivity are the only features which occur enough to maintain a strong consistency among the 26 data sets. They highlight that the other characteristics could apply to “small data” (Kitchin and McArdle, 2016, p. 8).

2. See the following for more information: http://www.narrativesinspaceandtime.ca/; https://intothedebrisfield.ca/; https://100years100stories.ca/; https://novascotia.ca/archives/virtual/default.asp?Search=THexp

3. In some ways, being a Francophone was helpful, yet also seen as suspect in regard to Francophone political history in Canada. As Québec developed its own national identity (being “Québécois” instead of French-Canadian), throughout the 1960s and 1970s, the discourse about the cultural and linguistic aspects of speaking French in North America was reframed in Québec under the larger question of seeking independent nationhood. Along the way, other Francophones in the country felt pushed aside or even betrayed. The Canadian Francophone reality outside Québec is still a blind spot in today’s Québec political and media scene.

4. Arendt’s conception of the politic is rooted in a separation of the public and private realms, which is at odds with the feminist claim that “the personal is political.” We strongly agree with these critiques and with Jones (2015) and Benhabib (1993) about how Arendt can still contribute to a feminist humanism through her optimism about political action.

References
Arendt, H. (1958). The human condition. Chicago, IL: The University of Chicago Press.
Arendt, H. (1961). What is freedom? In Between past and future: Six exercises in political thought (pp. 143–171). New York, NY: Viking Press.
Asberg, C., Thiele, K., & van der Tuin, I. (2015). Speculative before the turn: Reintroducing feminist materialist performativity. Cultural Studies Review, 21, 145–172. doi:10.5130/csr.v21i2.4324
Benhabib, S. (1993). Feminist theory and Hannah Arendt’s concept of public space. History of the Human Sciences, 6(2), 97–114. doi:10.1177/095269519300600205
Bivens, R. (2015). Under the hood: The software in your feminist approach. Feminist Media Studies, 15, 714–717. doi:10.1080/14680777.2015.1053717
Bowker, G. C. (2006). Memory practices in the sciences. Cambridge: The MIT Press.
boyd, d., & Crawford, K. (2011). Critical questions for big data. Information, Communication & Society, 13(5), 662–679. doi:10.1080/10919320.2012.678878
Crawford, K. (2016, June 25). Artificial intelligence’s white guy problem. The New York Times. Retrieved from http://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html
Data. (n.d.). Merriam-Webster dictionary. Retrieved from https://www.merriam-webster.com/dictionary/data
Dell’Aglio, D., Fruh, A., González Fuster, G., Kuzniar, N., Lim, Y. F., Lombard, A., & Wiedemann, K. (2017). International exploratory workshop on data ownership: Workshop summary. Zurich, Switzerland: Centre for Information Technology, Society, and Law, University of Zurich.
Denzin, N. K. (1997). Interpretive ethnography: Ethnographic practices for the 21st century. Thousand Oaks, CA: SAGE.
Edwards, R., & Mauthner, M. (2012). Ethics and feminist research: Theory and practice. In T. Miller, M. Birch, M. Mauthner, & J. Jessop (Eds.), Ethics in qualitative research (2nd ed., pp. 14–28). Thousand Oaks, CA: SAGE.
Geertz, C. (1973). Thick description. In The interpretation of cultures: Selected essays (pp. 3–30). New York, NY: Basic Books.
Geertz, C. (1998). Deep hanging out. The New York Review of Books, 45, 69–72.
George, E. (2014). Éléments d’une épistémologie critique en communication—Au carrefour et au-delà de l’école de Francfort, des cultural studies et de l’économie politique de la communication [Elements of a Critical Epistemology in Communication - At the Crossroads and Beyond the Frankfurt School, Cultural Studies and the Political Economy of Communication]. In F. Granjon & E. George (Eds.), Critique, sciences sociales et communication (1st ed., pp. 97–136). Paris, France: Editions Mare & Martin.
Gibson-Graham, J. K. (2015). Ethical economic and ecological engagements in real(ity) time: Experiments with living differently in the Anthropocene. Conjunctions: Transdisciplinary Journal of Cultural Participation, 2, 47–71. Retrieved from https://tidsskrift.dk/tcp/article/view/22270
Gillies, V., & Alldred, P. (2012). The ethics of intention: Research as a political tool. In T. Miller, M. Birch, M. Mauthner, & J. Jessop (Eds.), Ethics in qualitative research (2nd ed., pp. 43–60). Thousand Oaks, CA: SAGE.
Gilligan, C. (1982). In a different voice. Boston, MA: Harvard University Press.
Gitelman, L. (2013). “Raw data” is an oxymoron. Cambridge: The MIT Press.
Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. Feminist Studies, 14, 575–599. doi:10.2307/3178066
Hine, C. (2000). Virtual ethnography. Thousand Oaks, CA: SAGE. Retrieved from https://us.sagepub.com/en-us/nam/virtual-ethnography/book207267
Hoffmann, A. L. (2016, November). Toward a conception of data violence. Terms of privacy: Intimacies, exposures, exceptions. Presented at Terms of Privacy: Intimacies, exposures, exceptions. McGill University, Montreal, Québec, Canada.

Hoffmann, A. L. (2018). Data, technology, and gender: Thinking about (and from) trans lives. In J. C. Pitt & A. Shew (Eds.), Spaces for the future: A companion to philosophy of technology (pp. 3–28). New York, NY: Routledge.

Jones, K. B. (2015). Queer(y)ing Hannah Arendt, or what’s Hannah Arendt got to do with intersectionality? New Political Science, 37, 458–475. doi:10.1080/07393148.2015.1089024

Kitchin, R., & McArdle, G. (2016). What makes big data, big data? Exploring the ontological characteristics of 26 datasets. Big Data & Society, 3(1), 1–10. doi:10.1177/2053951716631130

Lutzko-Toth, G., Bonneau, C., & Millette, M. (2017). Small data, thick data: Thickening strategies in trace-based social media research. In A. Quan-Haase & L. Sloan (Eds.), SAGE handbook of social media research methods (pp. 199–214). Thousand Oaks, CA: SAGE.

Luka, M. E., & Lilley, B. (in press) The NiS+TS psychogeographer’s table: Countering the official Halifax explosion archive. Public 57: Archive/Anarchive/Counter-Archive. Toronto: Public Access in association with Intellect Ltd.

Luka, M. E., Millette, M., & Wallace, J. (2017). A feminist perspective on ethical digital methods. In M. Zimmer & K. Kinder-Kurland (Eds.), Internet research ethics for the social age: New cases and challenges (pp. 21–38). Bern, Switzerland: Peter Lang.

Lupton, D. (2015). Personal data practices in the age of lively data. Social Science Research Network. In: J. Daniels, K. Gregory (Eds.), Digital Sociologies. Bristol: Policy Press. doi:10.2139/ssrn.2636709

Markham, A. (2013). Underscoring ‘data’: A critical examination of a core term in scientific inquiry. First Monday, 18(10). doi:10.5210/fm.v18i10.4868

Markham, A., & Buchanan, E. (2012). Ethical decision-making and internet research: Recommendations from the AoIR ethics working committee (Version 2.0). Retrieved from http://aoir.org/reports/ethics2.pdf

Mauthner, M. (2012). Accounting for our part of the entangled webs we weave: Ethical and moral issues in digital data sharing. In T. Miller, M. Birch, M. Mauthner, & J. Jessop (Eds.), Ethics in qualitative research (2nd ed., pp. 157–175). Thousand Oaks, CA: SAGE.

McRobbie, A. (2015). Notes on the perfect. Australian Feminist Studies, 30, 3–20. doi:10.1080/08164649.2015.1011485

Miller, T., Birch, M., Mauthner, M., & Jessop, J. (2012). Ethics in qualitative research (2nd ed.). Thousand Oaks, CA: SAGE.

Millette, M. (2014). Contre-public et médias sociaux: le cas du collectif francophone TaGueule! au Canada anglais. In S. Proulx, J. L. Garcia, & L. Heaton, (Eds.), Contribution Dans L’univers Numérique. Pratiques Participatives À L’ère Du Capitalisme Informationnel [Counter-publics and social media: the case of the French collective TaGueule! in English Canada] (pp. 91–102). Québec City, Canada: Presses de l’Université du Québec.

Millette, M. (2015). L’usage des médias sociaux dans les luttes pour la visibilité: le cas des minorités francophones au Canada anglais [Social media use in the struggle for visibility: the case of French-speaking minorities in English Canada] (Doctoral thesis). Université du Québec à Montréal, Montreal, Québec, Canada.

Millette, M. (2016, June). TaGueule! A webzine that speaks out about politics, identity and culture in French Canada: When social media use is a form of political participation. Paper presented at Youth Political Participation: The Diverse Roads to Democracy, McGill University, Montreal, Québec, Canada.

Mulder, F., Ferguson, J., Groenewegen, P., Boersma, K., & Wolbers, J. (2016). Questioning big data: Crowdsourcing crisis data towards an inclusive humanitarian response. Big Data & Society, 3(2). doi:10.1177/2053951716626054

Poizat, J.-C. (2007). Interview with Étienne Tassin. Le Philosophosophe, 29, 11–40. doi:10.3917/phoir.029.0011

Proulx, S., Heaton, L., Choon, M. J. K., & Millette, M. (2011). Paradoxical empowerment of produsers in the context of informational capitalism. New Review of Hypermedia and Multimedia, 17, 9–29. Retrieved from http://www.tandfonline.com.proxy.bibliothèques.uqam.ca:2048/doi/full/10.1080/13614568.2011.552646

Roberts, S. T. (2016). Commercial content moderation: Digital laborers’ dirty work. In S. U. Noble & B. Tynes (Eds.), Intersectional Internet: Race, sex, class and culture online, 147–160. New York, NY: Peter Lang.

Wang, T. (2016, January 20). Why big data needs thick data. Ethnography Matters. Retrieved from https://medium.com/ethnography-matters/why-big-data-needs-thick-data-b4b3e75e3d7

Zefferino, A., & Hogan, M. (2015). Queered by the archives: No more potlucks and the activist potential of archival theory. In A. J. Jolivette (Ed.), Research justice: Methodologies for social change (pp. 43–55). Bristol, UK: Policy Press.

Author Biographies

Mary Elizabeth Luka (PhD, Concordia), is postdoctoral fellow at the University of Alberta and past Ban ting Fellow at York University. Her research interests include the modes and meanings of creativity and innovation in the digital age and investigating how civic, artist, scholarly, and business sectors are networked together.

Mélanie Millette (PhD, Université du Québec à Montréal [UQAM]), is an assistant professor of Communication—Social Media at the UQAM. Her research interests include the political and cultural uses of social media and specifically the issues of visibility and public participation among members of minority groups.