My Entire Life is Online:  
Informed Consent, Big Data, and Decolonial Knowledge

Anne O’Connell  
York University

Abstract
Research protocols require that informed consent is secured before a project is underway. In the collision of online and offline worlds, researchers continue to offer research participants forms of confidentiality and privacy that no longer exist. The explosion of big data and digital information has transpired with little attention to ethical considerations of consent, privacy, and confidentiality central to research ethics, as these online practices enter the terrain once the preserve of scholars. In this paper, I am concerned with the ethical, epistemological, and political implications of the reorganization of knowledge production in the age of new technologies, big data, and “informational capitalism” (Coll, 2014). Critical debates about research ethics must be harnessed to challenge the rigid, individualistic, and corporate view of consent and privacy now being fashioned by privacy experts. In particular, decolonial and Indigenous research protocols developed in response to settler colonialism exposes the dangers and possibilities of what lies ahead. As a new biopolitical force, check-box consent has become a form of conduct that elides the public–private entanglements of the Internet, new forms of surveillance/privacy, and knowledge(s) that favour austerity (for some), neo-liberalism, and securitization.

Keywords: research ethics, informed consent, big data, surveillance, privacy, race

I was waiting in line in a Staples department store when the checkout clerk asked for my postal code and email address as part of a promotion the store was conducting. As one of the many requests we refuse or often passively agree to, I said I’d rather not provide her with any personal information. After a brief discussion she stated that she didn’t mind handing over demographic information to anyone, stating she had nothing to hide, after all her entire life was online. In The Imperial Archive: Knowledge and the Fantasy of Empire, Richards (1993) examines how the British Empire collected and collated information as a way of ruling colonies that were too far away to control. In many ways, what resulted was a “paper empire” (p. 4). As the state and private industry develop their knowledge-gathering capacities, my concern today is with our digital empire, a growing empire that operates inside and outside of institutional rules and privacy legislation, in ways that expand state power and surveillance practices evident in the passing of Bill C-
In an information age that heralds empiricism and data points supposedly free of theory, I am concerned with how contemporary studies that reproduce colonial and imperial ways of thinking and ruling become more difficult to name.

The terrain of research and ethics is quickly shifting beneath our feet. In the age of big data, endless amounts of information that account for our moment-by-moment offline and online engagements are stored by and sold to various stakeholders. In workplaces, social workers have little input into the creation and management of databases they use in their daily practice, which may affect their own job security and the lives of the people they work with (Reamer, 2012, 2013). Serious questions about confidentiality and the ethics of social media find non-profit agencies, the private sector, and government scrambling to design protective e-policies while the demand for e-services (online counseling and peer support groups, blogs) continues to grow. In academia, university research ethics boards employ research protocols and informed consent forms that protect a form of confidentiality and anonymity that no longer exists. Government, corporate, and academic practices continue to overlap, as corporate interests and emerging technologies have increasing power and sway in university and community settings. It has become increasingly difficult to distinguish between conducting research and gathering data points. Both are governed by separate legislation and institutional norms, even as they impinge on one another’s territory. Research projects increasingly create and access online data sources or rely on the outcomes they produce, while the world of big data has edged its way into forms of knowledge production often the preserve of scholars, yet ignoring academic research protocols. The collision of worlds in a context of austerity, securitization, and neo-liberalism makes for some tricky times. The ethics of research has a set of ethical codes set out since the Second World War, including the Nuremberg Code; the Declaration of Helsinki; the Belmont Report; other agreements from the World Health Organization (WHO) and United Nations Educational, Scientific, and Cultural Organization (UNESCO); and Canada’s Tri-Council Policy Statement (CIHR, NSERC, and SSHRC), which governs all research with human participants in Canada. The ethics of information gathering has a more recent tradition in Federal and Provincial Privacy and Freedom of Information Acts (1983), the Personal Information and Electronic Documents Act (PIPEDA, 1994),

---

1 C-51 increases police capacity for preventative arrest without warrants, cracks down on digital and hard copy “terrorist propaganda”, makes encouraging or promoting a terrorist act its own criminal offence, increases sharing of personal information with government departments and gives CSIS new powers to interfere and disrupt plots. The broad definition of all terms sets dangerous precedents (see https://openparliament.ca/bills/41-2/C-51/).

2 I draw upon boyd & Crawford’s (2012) definition which states that it is the “cultural, technological, and scholarly phenomenon that rests on the interplay of technology (maximizing computation power and algorithmic accuracy to gather, analyze, link and compare large data sets); analysis (drawing on large data sets to identify patterns in order to makes economic, social, technical and legal claims); mythology (the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with an aura of truth, objectivity, and accuracy)” (p. 663).
and the Personal Health and Information Protection Act (2004). More recent legislation deals with issues of consent in the anti-spam legislation, online privacy, and surveillance and copyright concerns (C-30; C-13; S-4).

Many of our institutional and government responses to ethical concerns are instrumental approaches for the “how to” of consent, autonomy, and privacy. While online and digital research protocols are being introduced in relation to these issues, the foundational questions about the communities we want to live in persist. Lévinas (1969) urged us to be aware of our ethical responsibilities when we are faced with another’s suffering. What kind of ethics are we moving toward when our face-to-face encounters are computer mediated, while the body itself is increasingly compartmentalized, exteriorized, and commodified through a tissue economy that includes biobanks (O’Doherty et al., 2011; Sariola & Simpson, 2011), transplant tourism (Jaycox, 2012), the human genome project, and reproductive technologies (Haines, Taylor, & Turkmendag, 2012)? In an age of increased abstractions, how do categories of race and racisms appear less visible or hyper-visible, with little interrogation into the concrete ways they are formed and used to organize knowledge production and ways of ruling? In Fatal Invention: How Science, Politics and Big Business Re-create Race in the Twenty-First Century, Roberts (2011) described how the human genome project finally laid to rest any notion of race as a biological category, which nonetheless has spurred on studies devoted to race-based science and technology. What does it mean to say I have nothing to hide, my entire life is online, when a wide set of stakeholders watch, exchange, and transform both the material and virtual identities we inhabit and create in targeted ways that appear to be neutral? As Cool (2012) observed, we must be careful of how “technological determinism and disembodied subjectivity resurrects the liberal human subject” (p. 27). With the growth of web-based surveys, large data sets, and algorithms, how do we recognize, respond to, and care for one another in research?

In this paper, I explore how the critical debates on informed consent are pertinent to our increasing investment in online identities, data, and activities. As we know, our data exhaust (by-products of our online activities) is collected and traded by governments and corporate entities, the data barons of today. In the age of big data, ethics are often non-existent, muted, or occurring in the moment of engagement (or after). As we scramble to respond to the new questions raised by online interactions and research, there are decades of critical debates about research ethics that have something to offer the world of big data. In the first section of this paper I lay out some of these broad debates and discuss how critical and decolonial approaches may be useful to the digital world. A second focus of this paper is on how the rise of online and digital data and its corresponding legislation is actively shaping public and private worlds, privacy and surveillance, and what constitutes knowledge. More than 98% of the world’s information is now stored digitally, and the volume of that data has quadrupled since 2007. Mayer-Schönberger (CBC, 2013b) described this as the datafication of life, showing how one exabyte was filled from collecting data from the beginning of time to 2003. Yet in 2013, we gathered 5 exabytes in 2 days. Smolan (CBC, 2015) stated how Facebook has uploaded 50
billion photos, while Walmart conducts one million transactions per hour. The future of the Internet of things is already here, hoping to further integrate the connectivity of our devices, objects, and services, while robotics and artificial intelligence encourage the further integration of flesh and machine. The sheer amount of data and the new programs and algorithms to make sense of it (at such little cost; Sullivan, 2013), has created new forms of knowledge and new disciplines—such as behavioural advertising, massified research, computational social science, learning, and people analytics. These shifts in knowledge production expose new vulnerabilities in our understanding of informed consent and research ethics. Through an examination of scholarly and popular literature and recent legislation, I argue that consent practices must be situated and understood by their shaping forces. These include three major areas: the public–private entanglements of the Internet, the shaping of privacy as surveillance, and a critical examination of big data as knowledge. Some attention to the dominant forces defining and shaping our online and offline worlds may assist us in rethinking consent and ethical ways of being in our research practices.

Informed Consent: Western Regimes and Decolonial Research

The scholarly literature on research ethics and informed consent in the West point to the influence of Plato, the Hippocratic Oath, and to the series of enlightenment theorists who explored and expanded upon liberal tenets such as individual rationality, autonomy, freedom, choice, and the social contract (Manson & O’Neill, 2007; McStay, 2013; Reamer, 2012). While many scholars accept these principles as universal and shared standards, some point to a number of problematic assumptions in these frameworks. Others characterize informed consent as a tool embedded in the violence of universalized colonial epistemological regimes (Khan, 2005; Mignolo, 2009; Pateman, 1988). Indigenous scholars, in particular, have critiqued the colonial knowledge base of research and have put forward an Indigenous philosophical, epistemological, and methodological approach that upends Western research practices (Battiste, Bell, & Findlay, 2002; Kovach, 2009; Tuhiwai-Smith, 1999). In this view, alterations to informed consent will never dismantle the dominant power relations and colonial histories upon which it rests and through which it is authenticated. As Manson & O’Neill (2007) argued the “18th century Enlightenment tradition of the social contract and the principle of freely given consent lends moral legitimacy to actions which would otherwise be regarded as unacceptable” (p. 57). Informed consent is embedded in a Western social contract (or racial contract as Charles Mills would say) that recognizes (certain) individuals as autonomous and free beings who must fully understand and voluntarily choose to participate in research studies. As Indigenous scholars have argued, this Western, liberal, individualist, and universalizing approach of settler colonialism runs counter to Indigenous values of relationship, community, reciprocity, respect, experience, and storytelling (Kovach, 2009).
If we examine the common institutional origins of informed consent, the literature begins with the Nuremberg Code\(^3\) (1947), set in place in response to the horrors of human experimentation in Nazi Germany, followed by the Declaration of Helsinki\(^4\) by the World Health Association (1964), which outlined ethical principles for medical researchers and was revised for the seventh time in 2013. While the Nuremberg Code and Declaration of Helsinki are not legally binding, they have been adopted by national legislations and across many disciplines, associations, and professions, in and outside of clinical studies. The main tenets uphold the view that research participants must know what the research is and why it is being done, and must have the capacity to voluntarily consent to the parameters of the study without coercion. Another central document is the Belmont Report (1978), which articulated three core principles to consider when researching human subjects: respect for persons, beneficence, and justice. The Belmont Report arose out of the now-famous Tuskegee Syphilis Study (1932–1972), in which hundreds of rural and impoverished African American men, many of whom had contracted syphilis, were never made aware of or treated for their condition, so that the U.S. Public Health Service could study the progression of an untreated disease. A parallel study transpired in Guatemala (1945–1956), where orphans, inmates, sex workers, and psychiatric patients were infected with sexually transmitted diseases without their knowledge. An apology from Secretary of State Hillary Clinton in 2010 has been followed by two attempts at class action suits (Ellis, 2015). These studies come from a colonial history of human experimentation and ideas of racial superiority during racial slavery, the rise of racial sciences, and the eugenics movement (Goldberg, 1993, 2001; Hartman, 1997; McLaren, 1990; O’Connell, 2009, 2010a, 2010b). Another central player in the world of bioethics is the Council for International Organizations of Medical Sciences (CIOMS, established by WHO and UNESCO in 1949), who have produced more detailed and exacting standards, governance arrangements, and guidelines. In Canada, we have the Tri-Council Policy Statement of 1998 (CIHR, NSERC, and SSHRC) which governs all research with human participants under the principles of respect for persons, concern for welfare, and justice. Through these reports, Institutional Review Boards (United States) and Research Ethics Committees (Canada) have been established to review and approve research studies and informed consent procedures, often operationalized through the informed consent form. Each form is required to include adequate information about the risks and benefits of the research in ways that support comprehension and voluntariness.

---

\(^3\) Ghooi (2011) argued that the Nuremberg Code has no legal force and drew heavily from the Guidelines for Human Experimentation (1931) already well known in Germany, yet these earlier guidelines are never referenced by the mostly American authors of the Nuremberg Code.

\(^4\) Revisions in the past have outlined conditions for proxy consent, the establishment of Institutional Review Boards (IRB) and Research Ethics Committees, the importance of the consent of minors, and the problem with HIV trials using placebos in developing countries while patients in the U.S. have full access to drugs. More recent changes provide compensation to participants for research related injuries, and the dissemination of research results including unsuccessful studies.
An emphasis on informed consent and a review by a research ethics committee is an agreed-upon standard that is designed to reduce risk to participants and to protect their welfare.

The critiques of informed consent range far and wide, many of them applying quite directly to digital ethics. Violations and failures of informed consent processes can be glaring or more subtle, as they are often unable to respond to a variety of different research contexts and locations. Many argue that informed consent protocols reproduce a dominant medical model and a rigid view of power in clinical and non-clinical research. Universal standards never uncover the material and cultural inequities of research itself (i.e., North–South funding inequities), the problematic assumptions underlying research studies (Western colonial epistemes), and the institutional and organizational practices that configure the researcher and participant role (university vs. community). In the clinical arena we often neglect the social and institutional constraints on people’s decisions (Degner, 2002; Sinding & Wiernikowski, 2009), and recycle a vision of autonomy at odds with human interdependence (Sherwin, 1998). An extensive literature on research ethics in relation to medical research and HIV/AIDS clinical trials in “resource poor” countries explores the structural inequities, barriers, and complexities. The literature covers the dilemmas of working in countries with no research ethics boards, literacy and language barriers, and the use of proxy consent or oral consent. Researchers highlight issues with stigma, sampling, recruitment, and the ability to educate, build local capacity, community involvement and collaboration. Debates also revolve around compensation rates, data security, local politics, hierarchies and inequities, and levels of funding and sustainability, (Angwenyi et al., 2013; Bayer & Edington, 2009; CAHR, 2008; Fang, Steen, & Casadevall, 2013; McCreesh, Tarsh, Seeley, Katongole, & White, 2013; Mystakidou, Panagiotou, Katsarakis, Tsilika, & Parpa, 2009; Penn & Evans, 2010; Sariola & Simpson, 2011; Shah, 2012; Shamin & Qureshi, 2013).

Given the gross national and global inequities the underlie research projects, how can people and communities make informed choices (if that itself is possible) when the material preconditions for decision making are absent (Haimes et al., 2012; Jaycox, 2012)? Studies on the medical economy of transplant tourism and stem cell research reveal the ways in which impoverished women “consent” to selling their organs and participating in stem cell research. The issue of inequality and differential medical treatment was recently highlighted by the Ebola outbreak and who was to receive an untested experimental drug. The rise of medical colonialism is witnessed by the shift of biopharmaceutical clinical trials from the West to Brazil, China, and India—trials that test for diseases not even prevalent in these communities (Kamat, 2014). In cases in which structural inequalities are less severe, Hochhauser (1999) argued that research descriptions and consent forms are incomprehensible to many. Forms that outline research objectives and procedures are designed to protect the safety of institutions, operating as a legal contract with an assumed rational actor that in the end has little effect on a patient’s decision-making process (Armstrong, Dixon-Woods, Thomas, Rusk, & Tarrant, 2012).
Qualitative researchers question the dominance of the medical model when trying to attain informed consent. Ethnographers argue that consent is not something that happens prior to research but is ongoing and embedded in the research relationship. Informed consent prohibits the trusting and respectful relationships and forms of inquiry that social scientists prefer (Boulton & Parker, 2007; Murphy & Dingwall, 2007). The normative top-down expression of power interferes with this process by imposing a singular standard for consent that is based on the idea that the researcher always has more power and no risk compared to the participant (Nordentoft, Helle, & Kappel, 2011). The risk model imposed upon social scientists is unnecessary for studies that are equivalent to the risk of living everyday life. While scholars critique the imposed risk model, this position can equally advance the innocence of research and dismiss the risks associated with building intimate relationships and the everyday risks some people experience. Some scholars also claim that institutional barriers to approving one’s ethics reveals the over-protective paternalism of liberal individualism; while others are concerned they do not do enough to protect vulnerable groups (Barton, 2011; Clement & Bigby, 2013; Fisher, 2012; Shorey, Cornelius, & Bell, 2011; Swaine, Parish, Luken, & Atkins, 2011; Tylldum, 2012). These arguments tend to reproduce divides between individuals, caregivers, and communities along with disregarding the ways in which institutional power operates to apply extensive and subtle pressure on a participant’s ability to consent or not (Humphrey, 2013; Plankey-Videla, 2012; Shumaker & Medoff, 2013; Wolbransky, Goldstein, Giallella, & Heilbrun, 2013). These complicated and historically organized roles in institutions need to be considered by researchers who claim to advocate on behalf of the vulnerable by stressing more open consent models.

Much of the literature on research ethics and informed consent attempts to find room for improvement by reducing risks and encouraging recruitment, more participation, and individual and community benefits. These attempts to open up the process recognize how fully the research process is embedded in relations of gender, class, race, sexuality, and ability (Kendall & Rogers, 2007; Strom-Gottfried, 1998). Yet these claims are further complicated by critical race theorists, post-structural and governmentality theorists who recognize the productive power of informed consent. The idea is not to protect individual participants but to examine what kind of participant is being construed through these practices. In the name of improvement, openness, transparency, and accountability, increased pressure on participation and informed consent creates an individualized and responsibilized research subject and patient. Sariola and Simpson (2011) and Sinding and Miller (2011) noted how in health care in particular, a more participatory and open approach can exacerbate inequality. Power shifts from medical professionals to research participants who become the locus of decision making about research and their own health care, (Sariola & Simpson, 2011, p. 517) an endeavour that overburdens women and presupposes a set of resources that only those most privileged can access (Sinding & Miller, 2011). Reubi (2012, 2013) examined how shifts in the bioethics literature in Britain have moved from supposedly paternalistic approaches in the past to the formation of new ethical health care relationships characterized by dialogue and
respect and a subject who reflects and decides upon their own participation in research. As Reubi (2012) observed, “paternalism becomes rhetorical rather than historically accurate … its depiction in such a negative light opens up the space for intervention and reform” (p. 356). This latest configuration of the informed participant is a neo-liberal subject who is able to make informed choices to produce an individualized responsibility for health care itself. The normalization of these expectations and duties aligns with a consumer model of medicine in the name of public altruism, openness, protection, and progress in a neo-liberal context that rewards the individualization and privatization of health care from WHO to community-based agencies (see O’Connell, 2015; Sinding & Miller, 2011).

Further challenges to research ethics and informed consent come from Indigenous scholars and communities that turn away from incremental improvement and instead disrupt the material and epistemological foundations of research. In addition to articulating Indigenous philosophies, many Aboriginal communities and organizations in Canada are adopting a variety of Indigenous protocols, including the principles of OCAP (ownership, control, access, and protection) developed to govern the First Nations Regional Longitudinal Health Survey (First Nations Centre, 2007). Castellano (2004, 2010) noted how aboriginal ethics have a different set of ontological and epistemological starting points. An ethic of non-interference and an understanding knowledge based on connections between land, family, spirituality, values, and everyday living take precedent. The centrality of oral traditions, the knowledge of Elders, ceremony, and language place Aboriginal philosophies in opposition to positivist scientific research. Ethics reviews do not only apply to working with human subjects given the importance of ancestors, the land, non-human entities, visions, and a spiritual world. Dialogues between these entities are lost, and data are reduced to the observable phenomenon of human action. Aboriginal research and science strives for a “holistic awareness and highly focused analysis” whereby individual stories are validated by communities, unlike studies that call for objectivity and distance. Indigenous research is tied to self-determination and ownership and control over data (Castellano, 2004, pp. 102–104). Many of these principles are central to OCAP, in which relational accountability and respectful representation are key. Indigenous communities must maintain the intellectual property rights to own their knowledge, while ensuring the epistemological and ethical values of holism, spirituality, and interconnection are upheld. The Panel on Research Ethics (2010) recognized Aboriginal communities as a unique population and laid out key concepts and definitions, and ethical frameworks in Aboriginal contexts. Indigenous scholars draw on these ethical foundations, applying them more recently to digital data management, viewed as a site for decolonization and resurgence (McMahon, LaHache, & Whiteduck, 2015). Documenting a recent project in the Mohawk community of Kahnawâ:ke, these authors found that the ability to control community data in the areas of research, education, health, finance, membership, housing, land, and resources was central to self-determination. How might these critiques and alternatives contribute to the ways in which digital ethics are being construed? At the same time, how are the broader forces surrounding the digital world transforming research ethics protocols?
Digital Ethics: 
Public–Private Entanglements, Surveillance, Privacy, New Knowledge(s)

For the most part digital research and big data have sidestepped decades of deliberation about consent, while profiting from and privatizing data and attaching knowledge practices to forms of surveillance. Nonetheless, we continue to practise informed consent in our research; but can we deliver on what we promise? How can confidentiality and anonymity be practised in the same ways when they are being transformed or ignored in the digital world? What does it mean to rely on secondary data of individuals in the public domain, when clear protocols explaining the purpose of the research were never presented? How might our research findings be linked to identities or profiles tracked by Facebook or Google? More and more research is conducted using web-based surveys, collecting and linking data points and private information from health records, institutional and government records, phone and text records, tissue samples, public blogs, forums, and social networking sites as possible data. How might consent be viewed and operationalized by scholars located in university settings in relation to data collectors in government, non-profits, and commercial bodies? Research protocols advise that if the site is reasonably understood to be public, then consent is not required; yet the data should remain anonymous even when texts and identities are already public (Boggio, 2010; Foster & Young, 2012; Wager, 2012; Wilkinson & Thelwall, 2011). Opt-in policies, de-identification protections, and the introduction of privacy guards on data are policies that have been put in place. Yet larger questions persist. As the critical literature on informed consent shows, material inequities and challenging colonial knowledge formations need to be at the centre of our research. A biomedical model still dominates, while institutional concerns with privacy eclipse the deeper critiques of the colonial and liberal antecedents of informed consent protocols, proposed by Indigenous scholars and decolonial research (Mignolo, 2009; Tuck & Yang, 2012). In order to maintain a critical research agenda in online and offline worlds, I argue that we need to attend to and challenge the legislation and protocols that are now reshaping public–private entanglements, privacy and surveillance, and new knowledge forms.

Public–Private Entanglements

The collection and surveillance of our data continues to challenge and alter private versus public distinctions. When are we being private, when are we in public versus being public (boyd and Crawford, 2012)? As a problematic distinction that organizes modern liberal power itself (i.e., private = family, economy = public), the differences between government and corporate terrains (public–private) are becoming increasingly opaque in a wide range of areas such as health, security, surveillance, and the ownership of information. In spite of Indigenous scholars and communities who stress ownership and control of research, federal and provincial legislation that “compels First Nations to share information with third party organizations and the public” point to the ongoing effects of settler colonialism (McMahon et al., 2015, p. 4). Is information gathering done for marketing and
private business purposes, or is this done to further health research in communities? Is one nobler or more in the public interest compared to the other, when the public and private are being reconstituted? As Bennett, Haggerty, Lyon, & Steeves (2014) added, the conflation of these two realms is largely driven by the unquestioned principle of efficiency. Yet the kinds of consent being enacted, or not, in these terrains are critical to ethical research, professional practice, and our everyday lives. With such scale, Google can track flu outbreaks faster than public health departments, the collection of health data from premature babies can accurately predict an imminent infection, sensors built into carpets can be monitored by family at a distance and warn them of anomalous movements of an elderly person living on their own. A new app can predict the onset of depression (CBC, 2015). These overlaps are critical to understand. Especially to Hugo Campo, who attempted to access the data from his pacemaker to see if his health was improving yet was told by the manufacturer that this information belonged to the company (Salber, 2012). Or to Ellen Richardson when she tried to cross the U.S. border for a holiday recently but was turned away due to health information shared with border services about her suicide attempt two years earlier (Hauch, 2013). Like many others to whom this has happened, it is only through extensive public exposure that we have an awareness of this issue, with Privacy Commissioners now promising a full investigation after the fact. A class-action lawsuit was filed in June 2014, after 8,300 health data files of new mothers from Rouge Valley Hospital were sold to a financial firm (Marguson, 2014). Concerns about the mass of information being collected and exchanged between public and private entities will continue to fill our courtrooms and legislatures. We will continue to prosecute individuals for acts that mask the deeper problems with informed consent and the collection of online data in the first place.

Audrey Tobias was touching the tip of the (melting) iceberg when she refused to fill out the mandatory census because it was being processed by Lockheed Martin, a U.S. arms manufacturer. As a peace activist she questioned the choice of a weapons manufacturer to manage our data and the possible security breaches that could transpire (CBC, 2013a). In a recent mock exercise at the Canadian Department of Justice, 37% of staff clicked on a phony phishing link in their email, raising questions about the security of our information, as Tobias warned (CBC, 2014a). The Heartbleed virus that hacked into Canada’s Revenue Agency was the work of a 19-year-old computer science student (“Alleged Heartbleed hacker,” 2014). Bolen (2014) reminded us that 85% of Canada’s Internet traffic goes through the United States, while our government actively collaborates with the U.S. National Security Agency (NSA) in clandestine surveillance operations in 20 countries. When did “we” consent to all of this? As Gandy (2000, cited in McStay, 2013, p. 599) wrote, “while our mediated world becomes increasingly transparent, those who seek to profit from our data are incredibly opaque”. With new private–public entanglements how do we remedy these issues? As the research literature shows, structural inequalities continue to predetermine privacy, autonomy, and informed consent. Who has the financial power and expertise to collect and organize data, and how do individuals negotiate with opaque but real institutional forms of power that are public, private, or both? More than ever, informed consent practices need to
recognize these differences as key to ethical inquiries. How can individuals and communities have ongoing knowledge and ownership of and access to their own data? The notions of risk and benefit to being online remain unclear in privacy and information legislation and the Tri-Council protocols. These concerns become more difficult to define and examine as digital worlds are omnipresent, privatized, and commercialized outside of public bodies. The blurring between the public and private for users and in how the digital world itself is owned and organized complicates ideas of consent. Those who are without digital goods and services are seen to be outside of the economy, making the constant expansion and upgrade of digital goods and services essential to life itself. Colonial scripts of improvement and civilized conduct are alive and well in Zuckerberg’s (2014) recent statement that the entire “developing” world must be wired if they are to participate in the global economy and be pulled out of poverty.

**Shaping the Internet**

Discussions of preserving someone’s autonomy, confidentiality, and privacy are complicated by the public and private entanglement of the Internet itself and by legislative attempts to monitor and “protect” our data. As Foster and McChesney (2011) noted, the Internet would never have come into existence if it had been left to the private market. What started as a military and university endeavour has quickly devolved into the growth of monopolistic corporate powers that benefit from government deregulation and the elimination of competition. Calls for making broadband and the Internet essential public utilities are being closely debated, as the Federal government promises to extend digital access to rural communities, forgotten by the private sector in ways that still put Canada behind. The Competition Bureau of Canada continues to show how major players overcharge and dominate the telecom market, while the digital divide grows in Canada. Statistics Canada shows that 83% of Canadians use the Internet overall, but that number dwindles to 25% for low income households that access Internet wireless services (Geist, 2013). Foster and McChesney (2011) compared the digital divide in the United States to similar developments in health care, both allowing enormous corporate profits for the worst services. Debates over a two-tiered Internet (slow and fast lane) are met with calls for transforming it into a public utility (CBC, 2014b). In Canada, the commitment to net neutrality has morphed into a system that caps data usage, creating ceilings and requesting payment for more access.

Further areas of concentration are evident in search engines, digital download companies, and the Wi-Fi market. Mager (2012) tracked how search technology started out in the academic realm, yet became increasingly commercialized and based in consumer profiling to adjust advertisements to users’ individual interests (p. 771). Google’s move into smart phones increased its data points in order to build detailed user profiles, while alliances with competitors helped extend its position as default search engine to mobiles. Fuchs (2011, cited in Mager, 2012, p. 781) called Google the “ultimate user-exploitation machine” exploiting connections and networks of website providers and users’ activities, now dictating how website
providers must build their sites and punishing those that destabilize its system. It continues to control 75% of the market, making a mockery of economic and neoliberal notions of consumer power and cutthroat competition (Foster & McChesney, 2011). As Vaidhyanathan (2011) added, we are not the customer of Google, we are its product, as every site we visit creates a profile sold to advertisers. The absence of informed consent is particularly jarring given we are not only participants in this endeavour, but are the product and consumer. Similar to the Internet, Vaidhyanathan (2011) maintained, Google should be made a public utility, an interesting proposition given that the expense of their data sets and enormous archive are affordable to universities with the largest research budgets. As Foster and McChesney (2011) added, we need to challenge a world of “digital feudalism, whereby a handful of colossal corporate mega-giants rule private empires” (p. 4).

Shaping Privacy

Ideal notions of an open, free Internet outside of government regulation are increasingly countered by calls for fundamental digital rights protected by government. Yet, digital and individual rights are being trampled by governments in the name of security, as revealed by Edward Snowden and recent reports about the policing relationship between telecom companies and the federal government of Canada (Bolen, 2014). Displaying a circular logic, public–private actors must bypass consent and invade one’s security in the name of security. As Genel (2005) argued, the paradox of biopolitics is that protection for some is fully tied to harm for others; others who must be positioned as intolerable, outside of humanity. Colonial and imperial logics are built on knowledge practices designed to define and manage populations, and establish the right to rule. In Canada, new forms of surveillance in relation to data privacy have relied on three intolerable others: the pedophile, the cyberbully, and the terrorist. As we know, the handing over of our data from Canadian telecommunication companies to police and intelligence agencies is widespread. As recently reported, the federal government requests personal information from telecom companies 1.2 million times each year (only 3 out of 9 companies reported this). Canadian Border Services makes thousands of requests for customer data every year, without warrants. As Bolen (2014) argued, unlike the uproar of Edward Snowden and the NSA, the story in Canada was dead after a week, with many Canadians stating they have nothing to hide. To speak out and resist means you have something to hide. Many racialized and Indigenous scholars and communities (and retired Supreme Court Justices) have challenged the dangerous precedents set in the Anti-Terrorism Act, C-51, and noted how stirring anti-Muslim fervour is used to shape privacy and undermine protest. The idea that one has nothing to hide aligns with the privatization of policing and surveillance practices that have less oversight, enhancing states of exception in our everyday lives. Without the knowledge, permission, and consent of the user, telecom providers have become key players in data collection and in disciplining and punishing citizens. The ethical mainstays of research—informed consent, confidentiality, and comprehension— are absent as privacy legislation informs the collection of our data, outside of our full
knowledge in spite of those who challenge these measures (Dobby, 2014). How does this legislation affect our research?

As we have witnessed, the federal government has linked the digital world with increased surveillance measures, a policy shift that requires intolerable others. In the first instance, Bill C-30, Protecting Children from Internet Predators Act (withdrawn due to opposition) allowed for more surveillance of online accounts without warrants and relieved Internet providers who provided information from any legal responsibility. In a failed attempt to sell the bill, Federal Minister of Public Safety Vic Toews initially announced that either you are with the government or you are with the pedophiles (despite the fact the only mention of children was in the title of the bill). The follow-up Bill C-13, marketed this time as the anti-cyberbullying bill, offered full criminal and civil immunity to those handing over information and an expanded scope of access to information for public officials and private-sector organizations. Introduced in November 2013, Bill C-13 recommended lower thresholds to access online data in the name of extending anti-bullying measures (CTV, 2014). Carol Todd requested that the government stop using her child’s name when attempting to undermine privacy (Boutilier, 2014a). More recently, the Supreme Court through the Spencer decision supported the right to online privacy and “eviscerated the notion of voluntary disclosure” that Toew’s bill attempted to introduce (Geist, 2014b). Nonetheless, Bill C-13 was passed in October 2014, in spite of being at odds with the Spencer decision and widespread criticism (Wingrove, 2014). These increased powers of surveillance have spread to cellphones that, due to a recent Supreme Court 4-3 decision, now can be searched by police upon arrest (Payton, 2014). Canada’s Privacy Commissioner Daniel Therrien opposed sections of C-13, yet even his appointment raised concerns given his long career as a government lawyer responsible for national defense, public safety, and immigration policy (Boutilier, 2014b). Other scholars have noted how the dual and conflicting mandate of privacy commissions in the first place de-legitimates arms-length inquiries into government surveillance (Bennett et al., 2014). Yet these productive paradoxes endure in liberal democracies that become the arbiters of their own transgressions (see investigation into G-20 kettling in Toronto [Mandel, 2014] or the recent release of the CIA’s torture practices [Mayer, 2014]).

While the privacy of residents is under threat, their ability to access information from and about the government has diminished: The government’s compliance with freedom of information requests has stagnated or regressed as paper-based mailed-in forms remain the norm (Beeby, 2013). While the origins and definitions of privacy certainly vary, Coll (2014) asked us to be aware of how privacy is being shaped and by whom. Heeney (2012) suggested that we need a “reconceptualization of privacy and protection from information entrepreneurs and omnibus information providers” (p. 316). This request, however, hands over fundamental and enduring discussions about privacy, autonomy, self-determination, and community to the private sector. Is Rogers Communications Inc. the appropriate body to be establishing privacy policies and protocols outside of public scrutiny yet accountable to shareholders? Instead of viewing privacy in opposition to various

Intersectionalities (2016), Vol. 5, No. 1
Special Issue: The Ethics and Politics of Knowledge Production
forms of surveillance, Coll (2014) suggested we see how they act in concert with one another. Coll (2014) argued that the idea of privacy has been taken over and “re-shaped by and in favour of informational capitalism, notably by being over-individuated through the self-determination principle” (p. 1). It produces “subjects of privacy who are supposed to take care of it according to the official conception of privacy advocates and of the legislature” (p. 1). Once it is defined, privacy becomes only about data and remains the right and responsibility of every individual instead of a collective value. As a form of conduct and self-conduct, users (and researchers) must demonstrate proper Internet conduct over themselves and their children. New privacy experts take over, calling upon the education of the masses in relation to managing their own privacy and that of their children. As Coll (2014) maintained, “private companies and governments seem to be actually defending privacy values much more than the majority of consumers,” although they are the very ones that represent a threat to privacy (p. 8). The idea that we have control over our data is chimerical given the highly secretive forms of data collection private companies engage in and own. Attempts to combat this inevitability are found in the OCAP principles adopted by Indigenous scholars, which continue to stress control and ownership of data in the context of hyper-individualism and settler colonialism.

Shaping Consent

While telecom surveillance and data sharing has occurred without users’ consent, other attempts at data protection in commercial and digital research realms include online consent, de-identification measures, privacy guards, and opt-in measures (Cavoukian, 2013). Recent legislation in England now requires advertisers who insert cookies on the websites users browse in order to tailor advertisements back to them, to obtain tacit consent through an opt-in clause (McStay, 2013). Similarly, Canada’s anti-spam legislation (CASL) for commercial messaging requires opt-in consent, clear identification and contact information about the sender, and unsubscribe options, that lasts for at least 60 days (Geist, 2014a). Informed consent issues are particularly important in relation to the use of web-surveys, big data, and secondary data in all sorts of sectors, such as health, education (learning analytics), and the labour market. In relation to genomic privacy with biobanks, electronic health records, and secondary data release, voices critical of consent practices are silenced by dominant views that all health research benefits society (Foster & Young, 2012). The use of learning analytics that require opt-in (checkbox) consent or that bypass consent altogether harvests student’s educational data with posts from Facebook and Twitter to optimize resources and improve programs, student advising, and decision-making processes. Some scholars posit that perhaps the panopticon can be replaced by a form of participatory or peer surveillance, as the

---

5 The former Information and Privacy Commissioner of Ontario, Ann Cavoukian (2013), described how the de-identification of data is a principle now around the globe. She argued that the risk of data violations is small compared to the usefulness of secondary research. Privacy guards can provide differential treatment of information so that more distortion can be added when data are more intrusive.
many watch the few, making consent irrelevant\(^6\) (Slade & Prinsloo, 2013; Sullivan, 2013). Again, technical solutions prevail as data are de-identified, given a short lifespan, and stripped of demographic information, and expert technicians such as data stewards and custodians are appointed in organizations (Cavoukian, 2013). Peck (2013, p. 81) described how beneficial big data can be when it comes to being hired, fired, retained, or promoted. Bias in hiring practices can supposedly be reduced when relying on web habits and may be more just measures of leadership, creativity, and decision making. Big employers of hourly workers such as call centres, along with Xerox and Microsoft, are turning to online assessments and statistical models that can pinpoint good trainers, and develop indicators that predict the entire life cycle of a worker.

Concerns about big data sets become a question of data protection, not a question about the ethics of the research question. As a less direct form of data collection, issues of harm and confidentiality appear less critical or are viewed as being already in the public domain. A moral script regarding the security, health, and knowledge of efficient institutions that participate in surveillance practices take precedence over issues of justice and the scrutiny of institutions. As mentioned earlier, patients and research participants become subjects who reflect and decide upon their own participation through check-box consent that encourages less dialogue about the research question. Similar to informed consent forms, online consent relies on an impenetrable and performative legalese that is reproducible, quick, and on an enormous scale can invite or bar participants from commerce, human services, or research projects while protecting institutions. Instead of examining the ways in which opt-in conduct becomes a marker of an informed digital citizen, these powerful legal and technical responses just need improvement. The danger lies not in participation, but in the rejection of participation. This was borne out in a recent court case in Europe in which a resident had to legally win the right to be forgotten online (Gollom, 2014; Mayer-Schönberger, 2014). Again, fundamental issues that critical and decolonial scholars stress in research ethics are undermined by the rise of research as informational capitalism. At the same time the health, education, and labour market systems are increasingly vulnerable, insecure, and subject to a differentiated, managerial, and effective consumer model (Brophy, 2012).

**Shaping Knowledge**

Many critics have raised concerns about the ways in which data points and algorithms are viewed as knowledge. As Bennett et al. (2014) argued, big data leads to a form of social sorting, as individuals become profiles sorted into hierarchies, whereby “certain kinds of profiles pass with greater ease than others” (p. 6). The

---

\(^6\) Users or participants are often aware that each online tool has its own structure for participation in relation to anonymity, pseudonymity, and identity play. Formats encourage local or global participants that use text or images in open or closed formats that delete or archive participation (Whitehead & Wesch, 2012).
surveillance of individuals has become a dominant organizational practice to categorize and treat people differently. Certainly, many recognize that “algorithms used by institutions invariably reflect and perpetuate current biases and prejudices” (Slade & Prinsloo, 2013, p. 1517). Harvested data are limited by time, and demographic data should not be attached to predict one’s chances at success (Slade & Prinsloo, 2013). Yet many scholars note deeper dangers (boyd & Crawford, 2012; Heeney, 2012; Kitchin, 2014; Mayer-Schönberger, 2014). Even with the removal of personal information and de-identification practices, people become a set of categories that injure their autonomy in addition to propagating views about individuals and communities. As many theorists note, Indigenous peoples and racialized and marginalized individuals cannot transcend the classificatory hierarchies of the communities in which they are deemed to be located. As Wiegman (1999) wrote, white subjects are afforded a kind of mobility and universality that never places them as a representative of a group. This becomes more and more dangerous as big data sets can produce “unbiased” correlations about groups we were unable to see in the past. As Mager (2012) noted, the sanctity of the individual exists even as it reduces the subject to an algorithmically derived profile. However, the scale of these data projects matched by their apparent ideological impartiality is deeply concerning. As boyd & Crawford (2012) described, big data ushers in a profound change at the levels of epistemology and ethics. While it certainly entails unprecedented levels of data collection and analysis, the idea that numbers speak for themselves glides over the reality that claims to objectivity and accuracy are misleading. These views re-inscribe the sometimes false divisions between qualitative and quantitative research, a genealogy so wonderfully mapped out by Poovey (1998) in The History of the Modern Fact. Data cleaning and errors, restrictive models that exclude “extraneous” data points, along with ideas of randomness and representativeness are compounded by research that shows people often have multiple online identities, insert inaccurate information, and differentiate between being in public and being public (boyd & Crawford, 2012, p. 672).

Kitchin (2014) described how big data inverts in some intriguing ways the research process we rely on. He noted how big data sets are the “by-product of another activity” and do not arise from a specific research question. A series of algorithms can be applied to arrive at the best explanatory model, unlike researchers who choose a methodology or analytic approach to apply to the data based on their own knowledge of the topic (p. 2; p. 5–7). Rather than testing a theory, Kitchin argued, “new data analytics seek to gain insights ‘born from the data’” (p. 4). As many have argued, this level of empiricism can present relationships and patterns that we do not know to ask. And while it can produce correlations, it cannot explain why—why things happen. Yet, others argue that correlation is enough. A research paradigm data rich and free of theory may be ushering in a new research paradigm. But as Kitchin (2014) added, all classifications, histories, data points, and disciplines are discursively formed, never exhaustive, imbued with values, perspective, context, and histories, some of which are favoured and others of which are erased (see O’Connell, 2009, 2010a). Without doubt, big data is more difficult to integrate into
the humanities and social sciences, in which context, politics, and history take precedence in a rich and competing theoretical terrain; yet it is not impossible.

The use of big data in social sciences runs the risk of producing studies that result in an inescapable conflation between correlation and causation. People of certain races do a particular activity more than others, so the activity becomes something they do because they are of a certain race. As Mayer-Schönberger noted (CBC, 2015), this leads to punishing people for things they might do, or suspecting someone of a crime because of who their friends are on Facebook. This form of racial policing is already well established in Toronto, through a carding system and subsequent massive database that targets young black men. Categories of race, gender, class, disability, and sexuality become fixed leaving little room for interrogating the forms of power that create and compose them and apply value to them in the first place. These epistemological concerns and biases are compounded by data that are owned by social media companies that restrict access to companies and universities. Only some of these entities can pay the market price for “full” or somewhat full data sets (boyd & Crawford, 2012). Others note how our knowledge systems are already restricted rather than expanded by the way search engines are organized. As many note, Google has become a proxy for cognition itself; one that sees the top ten sites accounting for 75% of page views (Foster & McChesney, 2011; Vaidyanathan, 2011). As Foster and McChesney (2011) stated “big sucks the traffic out of small” (p. 4). Even the digitization of archives has been unequal, as the documents and events of the colonized are often overlooked or omitted from data sets (Koh, 2014). Sullivan (2013) described the huge amount of institutional and corporate power in “an anti-democratic system of control that cannot be transformed because it can serve no other purpose than that for which it was designed—the rationalization and control of human existence” (p. 227). Reproducing and sustaining colonial and neo-colonial systems of domination will make human existence and suffering reliant on the assumptions about who we are and how we want to be and respond to the other. In thinking through research ethics and informed consent, it is critical we attend to the ways in which knowledge and being is reconstituted in the age of measurement.

Conclusion

Critical scholars have turned their gaze upon the Western imperial canons that dominate forms of knowledge and disciplines by calling attention to the problematics of research itself. The unequal relationships that surround informed consent reveal how material inequities and colonial knowledge forms need to be challenged. We need to do research in ways that allow for transparency, openness, and participation while remaining critical of these terms and how they operate in service to, and can enhance, neo-liberal goals. While governments and institutions scramble to create privacy laws, it is critical that we understand and challenge these laws and how they may shape our notions of privacy in research. Researchers must examine how new categories and forms of knowledge creation through surveillance measures are being commercialized and privatized in the name of security for all. These have serious
implications for the forms of privacy, autonomy, and relations that we promise in our research, ethical commitments that the digital world needs to consider. Whitehead and Wesch (2012) asked if digital worlds are part of reconfiguring what it means to be human. Given the proliferation of online worlds, gaming, multiple role plays and identities, and the further integration of flesh and machine, we need to trace how these worlds are made available to and shape particular subjects. Many are concerned about the ways in which online presences weaken our moral and ethical responsibility to one another. Certainly examples of social isolation, anti-social behavior, online suicides, bullying, trolling, cyberbullying, and cyberstalking are plentiful, and studies into them critical. But my concern is with how these problems about individual behavior are being put to use, and the kinds of individuals, subjects, and profiles (pedophile, cyberstalker, terrorist) that are being deployed to secure monopolistic corporate imperial power. Tufecki (2012) argued that the “human is always a contingent category and different regimes of ‘humanity’ have been deployed throughout history to produce the exclusions and inclusions so necessary for the construction of power through difference” (p. 4). Perhaps it is not so much the different identities but the different ways that identities, ethics, and opt-in conduct in research are now put to use for the privatization, commercialization, and securitization of the public.

References


Panel on Research Ethics. (2010). Research involving the First Nations, Inuit and Métis peoples of Canada. In Tri-council policy statement: Ethical conduct for research involving humans, 2nd ed. (TCPS2; Chapter 9). Ottawa, ON: Canadian Institutes of Health Research (CIHR), Natural Sciences and Engineering Research Council of Canada (NSERC), & Social Sciences and Humanities Research Council of Canada (SSHRC). Retrieved from


Tyldum, G. (2012). Ethics or access? Balancing informed consent against the application of institutional, economic or emotional pressures in recruiting


**Author Note**

Correspondence concerning this article should be addressed to Anne O'Connell, School of Social Work, York University, 4700 Keele Street, Toronto, ON, M3J 1P3, Canada. Email: aoconnel@yorku.ca