Social Media, Big Data, and Privacy:
Protecting Citizen Rights in the Age of Connection

Submission to the Committee on Access to Information, Privacy and Ethics

December 13, 2012

B.C. Freedom of Information and Privacy Association
103-1093 West Broadway
Vancouver, British Columbia, V5N 1E2
Phone: 604-739-9788 | Fax: 604-739-9148
fipa@fipa.bc.ca

FIPA would like to acknowledge the Law Foundation of British Columbia. Their ongoing support of our work in the areas of law reform, research, and education makes submissions like this possible.

INTRODUCTION

FIPA is a non-partisan, non-profit society that was established in 1991 to promote and defend freedom of information and privacy rights in Canada. Our goal is to empower citizens by increasing their access to information and their control over their own personal information. We serve a wide variety of individuals and organizations through programs of public education, public assistance, research, and law reform.

While our work focuses primarily on access and information rights in British Columbia, we have also played an active role in federal sphere. In 2009, for example, we prepared a submission for this committee focused on former Information Commissioner Marleau's twelve proposals for reform of the *Access to Information Act*. We have also appeared before you on Open Government in 2011, and we have been involved since the last century in exploring the privacy implications of various 'lawful access' proposals and other federal privacy issues.

The question of privacy in the age of social media is becoming extremely relevant to our work at FIPA, but more importantly, to the efforts of policy makers such as yourselves and to the everyday lives of Canadians. As more and more of our personal information circulates through complex global communication networks via social media platforms and apps, our conceptions of what privacy is, how it is safeguarded, and how it operates are challenged. It is important to stress, however, that contrary to what many social media evangelists might claim, such challenges have not diminished the democratic, personal, and social value of privacy.

Quite the opposite: as our lives become more deeply integrated with our social networking technologies, the costs of non-consensual privacy invasions dramatically increase. It is no overstatement to claim that a person's livelihood, safety and well being are now powerfully connected to the security of their information. As such, the need for a legislative and policy framework capable of addressing emerging conceptions of privacy is acute, particularly as laws such as the *Personal Information Protection and Protection of Electronic Documents Act* come up for amendment via Bill C-12. This committee's study of this subject is a promising first step toward just such a framework, and we are pleased to contribute.

A NEW PRIVACY FOR A NEW PUBLIC

Social Networking, Privacy, and the Limits of "Stranger Danger"

The claim that online social spaces threaten privacy is well rehearsed. Since the early days of the Internet, many have (correctly) argued that digital social networks are, at least in some measure, inherently privacy-invasive, and in many cases support dangerous, predatory behaviours. For this reason, social networking environments--including early entrants like LiveJournal, MySpace, and Nexopia, as well as more recent ventures like Facebook, Twitter, Google+, LinkedIn and Tumblr--remain powerfully linked, both in specific policy contexts and broader cultural frameworks, to what has been called a "stranger danger" discourse (Poyntz, forthcoming).

In this framework, social networks are seen as full of potentially predatory social actors who exploit our personal information for any number of nefarious purposes--extortion, blackmail, identity theft, bullying, and the like. That is, the invasion of privacy is taken to be an individualized and somewhat random problem, carried out unpredictably by mysterious figures that are difficult if not impossible to trace. As Livingstone and Helsper (2011) assert, this framing has led academics, popular critics, and policy makers alike to address the question of privacy in social networking, particularly where youth are concerned, by focusing on the regulation of individual behaviours through acts of blockage, restriction, and denial of access.

This is why, as social networking services have exploded in popularity over the past decade, we have seen a flood of best practices guides, pamphlets, policies, and education programs that emphasize strategies of non-disclosure. Kids and teens, for example, are often told that when it comes to protecting their privacy online, the best strategy is to opt out of such services altogether; to disengage from those environments that prioritize publicness through acts of self-disclosure. Certainly, these are important strategies with real merit that address how our online personas can be, and often are, exploited in damaging ways. The November, 2012 suicide of Amanda Todd, for example--a Vancouver-area teen who was mercilessly harassed through social media--is a tragic reminder that there is indeed a significant risk of exploitation involved in the disclosure of personal information through social networking platforms.

That said, when it comes to developing policies that aim to safeguard citizen privacy in social networking environments, this trend toward simply encouraging non-disclosure and individual opt out tactics is inadequate for two reasons:

1. It underestimates just how comprehensively intertwined our social, political, cultural, civic, and economic lives are with digital networks. Especially for young people, online social spaces, at their best, can function as important sites of social and informal learning, and may actually help kids develop into accommodating, democratically-engaged subjects (Poyntz, forthcoming). Further, if Canada is to become a leader in the new information economy, we must be willing to engage with the networks that support the circulation of innovative ideas. It does not seem realistic on a number of levels to simply disengage from potentially rich arenas of exchange and value creation. As award-winning journalist Emma Teitel put it succinctly in *Maclean's* in November, 2011, "the only way to avoid starring in party pictures on Facebook is to skip the party. And who wants to do that?"

2. By configuring privacy-invasive behaviours as largely individual acts, and by repeatedly linking those acts to the shadowy figure of the cybercriminal, it fails to account for the much broader, (infra)structural ways in which personal information is collected, used and disclosed in social networking environments, and completely overlooks the engine that in fact drives online social networking: **big data.**

Big Data and Metadata: How Identification and "Personal Information" are Changing

> As Jay Stanley (2012), Senior Policy Analyst at the Speech, Privacy and Technology Project of the American Civil Liberties Union, puts it, **big data** refers to how exponential increases in computing power, combined with the global expansion of digital communication networks and the plummeting cost of data storage, today make it possible to collect and store truly mind-boggling volumes of information. Once captured, this information can be analyzed in a process known as **data mining,** which generally refers to the use of automated processes and systems (like algorithms) to sift through huge volumes of data so as to "discover subtle patterns, correlations, or relationships" within a set (Stanley, 2012). Data mining is a way of "making things visible to us that have never been visible before" (ibid) and as Bigus (1996) and Cavoukian (1998) have written, is often done with commercial interests in mind. While exciting from a research and business perspective, data collection and analysis on this massive scale also raises serious concerns about the security and privacy of our personal information, and empowers distressing new forms of surveillance that in Stanley's words, "can further tilt the playing field toward big institutions and away from individuals."

But where does all this data come from? What activity or activities could possibly produce enough information to drive and sustain such immense collection and storage efforts? Outside of a simple increase in the volume of information shared by Internet users, one major factor contributing to the rise of big data is the tremendous expansion and intensification of video and audio surveillance over the past thirty years--both institutional surveillance in the form of security infrastructure, and informal or voluntary surveillance in such forms as mobile phone cameras and social photography networks (Instagram, etc.). But perhaps the most important driver of big data practices, outstripping the influence of even these major contributors, is not data itself, but the data *about* the data, more commonly referred to as **metadata.**

Every time we interact with a social media environment--even if that interaction is minute, mundane, or otherwise limited with respect to content--the environment itself, supported by complex data-gathering algorithms, produces huge volumes of information *about* the interaction: where it took place, when it occurred, what other users were around when it happened, what kind of device it took place on, what operating system that device was running, how long the interaction lasted, and the like. When these pieces of information are connected together, they can paint an extremely accurate portrait of the user, their preferences, their networks, and their habits. As Stanley succinctly puts it: "when you combine someone's personal information with vast external data sets, you create new facts about that person." That is, even in situations where users *don't* willingly share accurate, complete, or factual information about themselves, social media systems still collect such vast quantities of *metadata* that users become, in many ways, identifiable all the same. In social media environments, then, identification and identifiability is

as closely linked with interactivity as it is with conventional notions of disclosure. As Privacy Commissioner Jennifer Stoddard put it in a recent address to ETHI:

> "Social media companies can quickly amass a staggering amount of personal information…In addition to the preferences, habits and social interactions of their users, these companies also collect vast amounts of background information that are not visible on public profiles, including search histories, purchases, internet sites visited and the content of private messages. This collection of billions of data points allows social media companies — using sophisticated algorithms — to analyze user behaviour in order to refine their services and to identify ways to generate revenue."

In the world of big data, metadata, and social media networking, the "stranger danger" paradigm--the notion that online privacy ultimately rests on individual choices about what to disclose and withhold--is incomplete. Fake names, made-up birth dates, and imaginary hometowns do little to disrupt the capacity of social media platforms to identify users by linking vast metadata sets together in particular ways. A short way of putting this would be to say: if it is practices of interactivity that ground the privacy-invasiveness of social media platforms (ie. their ability to identify users), and if social media platforms are themselves built to facilitate those very practices of interactivity, then all social media platforms have a built-in element of privacy invasiveness.

But as mentioned, to simply disengage from these platforms could have negative effects on our economy, our social connectivity, and even our democracy. So how are we to proceed? How might we understand and practice privacy in a way that preserves the benefits of connection while recognizing and responding to the fundamentally new ways in which Canadian citizens are being identified--not just by predatory cybercriminals, but by algorithms, code, and metadata?

Contextual Privacy: Recognizing the 'Social' in Social Media

If metadata makes it possible to identify users based on how they interact with one another in social media environments--that is, if the possibility of identification and privacy invasion rests on the *sociality* of the system--then our notion of privacy, and the policies that protect it, must also be rooted in the social. As Valerie Steeves (2009), Professor of Criminology at the University of Ottawa has suggested, the explosion of privacy-related crimes like identity theft in the era of social networking is not simply a function of poorly implemented privacy policy. Rather, it points toward the fact that "we rely upon a definition of privacy that is problematic because it *strips privacy out of its social context*" (p. 193). For Steeves, privacy in a contemporary context exceeds "narrow procedural considerations of data protection;" the very considerations emphasized by the stranger danger paradigm.

In a blog post for *Public Policy and Governance Review,* for instance, co-Editor in Chief Max Greenwald recounts a speech given at the 2012 National Magazine awards by Emma Teitel, whose recent *Maclean's* essay on social media privacy "The New Paparazzi" is cited above. Teitel began her speech by reading aloud bits of personal information about some of the people in attendance--all of which she'd found publicly available on their Facebook pages. Even though all of the information Teitel read had been willingly disclosed by the audience members, the

experiment still produced feelings of privacy invasiveness, suggesting that acts of disclosure and the perception of invasion are more complex, flexible, and contextually-situated than a simple public-private split can grasp.

This is why privacy, in Steeves' estimation, must be understood and practiced as something that safeguards our friendships, kinship ties, and social lives *in general* against exploitation, misappropriation and invasion. If the information used to breach our privacy rights increasingly comes from our encounters in digital space, and not simply our *individual* patterns of disclosure, then any policy framework that would defend privacy must take seriously the concepts of interaction and sociality. It must approach the notion of privacy on broad, flexible terms so as to "question--and limit--the negative impact of surveillance on our social and democratic relationships" (p. 193) Steeves summarizes this position effectively by turning to Priscilla Regan's 1995 book, *Regulating Privacy:*

> "Privacy is more than an individual right; it is also a social good in and of itself that 'serves other [social] functions beyond those to the particular individual…' If the social value of privacy is not taken into account by policymakers, privacy will continue to shrink in the face of competing claims of security and convenience" (p. 194).

If, as Stanley writes, big data has the potential to shift power away from citizens toward already authoritative institutions, and does so precisely by mining and capitalizing on the metadata produced by our digital social lives, then our privacy policies must meet and counter this growing asymmetry on its own terms. This means departing from and accounting for the social, the interactive, and the shared.

POLICY CONSIDERATIONS

FIPA suggests that the Committee consider the following points in examining the issues surrounding big data, policy, and privacy in the contemporary Canadian context.

1. **Revisit and expand the definition of personal information in *PIPEDA***

   PIPEDA currently defines personal information as "information about an identifiable individual," not including "the name, title or business address or telephone number of an employee of an organization." This definition overlooks the fact that, as outlined above, the identifiability of a particular social media user is not necessarily contingent upon the collection of information "about" that user in any traditional sense. Rather, even ostensibly anonymous information such as algorithmically collected metadata now works to identify individuals.

   Further, as it is currently written, the definition presupposes an "identifiable individual," and so can only place restrictions on data collection that occurs *after* some earlier moment of identification. The moment of identification itself, then, necessarily falls outside the coverage of the *Act*. This configuration fails to consider that we are in fact *made identifiable* by engaging in discussion and connection with other users. That is, it is

precisely the function of social media services to identify us in the first place; as businesses and systems, they cannot operate without a preliminary (often extensive) identification of users through the collection of metadata. This act of identification is thus integral to the business transaction that occurs between social media users and proprietors, and so should fall within the protections of the *Act*.

Drawing on the contextual privacy framework above, privacy policy in the era of social media shouldn't simply protect information about identifiable individuals. Rather, it should also protect the social processes, practices, and interactions that *make us identifiable*, particularly where identification is necessary for functionality.
As such, FIPA suggests modifying and broadening the definition of what constitutes "personal information" to include the processes and practices of interactivity by which social media users become identifiable.

2. **Move against the 'voluntary sharing' regime that would be established by Bill C-12's amendments to *PIPEDA*.**

   This reconsideration of what constitutes "personal information" in *PIPEDA* should also come with a strenuous opposition to the privacy-invasive provisions laid out in Bill C-12. As Tamir Israel of the Canadian Internet Policy and Public Interest Clinic writes, C-12 would establish a "voluntary sharing regime" that avoids "even the most rudimentary of oversight and tracking typically associated with police surveillance." By broadening the range of entities capable of accessing personal information held by private sector organizations (including social media companies) in the name of "policing services," C-12 would erode existing statutory privacy protections. It would also dramatically extend the surveillance capabilities of vaguely defined 'lawful authorities' at a time when, to again pull upon the words of Jay Stanley, big data is already amplifying power imbalances between citizens and institutions.

   Importantly C-12 also moves us further away from a policy framework that acknowledges privacy as a social value that emerges through our interactions and connections. By expanding police access to personal information held, carried, and transmitted by telecom companies, social media enterprises, and Internet service providers, C-12 undermines the security of our practices of interactivity and sociality, and subjects digital social environments to intensified practices of surveillance. Coupled with the fact that even rudimentary, non-incriminating interactions between social media users can be implicated in unrelated suspicious activity as a result of geolocation and proximity metadata, C-12's voluntary disclosure regime threatens to wrongfully draw innocent Canadians into potentially criminal investigations.

3. **Bring Canadian political parties under federal privacy legislation**

   As social media networks have grown in popularity, sophistication, and scope, they have also become important sites of political mobilization. Shareable petitions and other forms of so-called "clicktivism" are extremely common today, and often spread virally through social media networks, collecting huge volumes of personal information as they do. Such petitions are often delivered to politicians and heads of state to pressure them to take action on key public interest issues.

In Canada, however, such actions carry a unique and surprising privacy threat, in that Canadian political parties are *not covered by federal privacy legislation*. This means that large data sets such as petition lists and even the voter registry can be used and disclosed in any way a party sees fit, free of all accountability or traceability. In this sense, Canadian political parties are complete anomalies. Every other organization in the country, after all (whether public, private, non-profit, for-profit, professional, or volunteer run) is mandated to protect constituent privacy to some degree, often under both federal and provincial legislation. FIPA sees no reason why political parties—arguably some of the most influential organizations in the country, not to mention some of the most seriously invested in social media networking and online community building—should be any exception to this rule.

The potential consequences of this loophole are significant. In the first case, it makes Canadians into targets for all kinds of exploitative, privacy-invasive behaviours, including identity theft, solicitation, and direct-call marketing. In 2009, for example, it was discovered that a Toronto cell of the Tamil Tigers, officially classified as a terrorist organization, had obtained copies of the Canadian voter list and were using it for fundraising purposes. Even cabinet ministers and major government figures have used petition data to promote partisan projects among particular interest communities. In the fall of 2012, for example, a number of gay and lesbian Canadians who had signed a petition to block the deportation of gay Nicaraguan artist Alvaro Orozco received an unsolicited email from the office of Immigration Minister Jason Kenney, promoting the Conservatives' efforts at protecting LGBT*-identified immigrants. This despite the fact that many recipients of the message had never voted for, supported, or otherwise engaged with the Conservative Party in a way that would require the disclosure of their sexual orientation or contact information. Rather, their personal data had been mined from the Orozco petition. Such unsolicited contact is made possible only by the unaccountable and untraceable sharing of citizen information within and by political parties not covered by privacy law.

Secondly, it could have serious effects on voter confidence, presenting a threat to political participation. As our political lives become increasingly entangled with digital environments that collect huge volumes of data on our choices and habits, it is essential that government officials, campaign volunteers, and party staff take steps to ensure the security of that data—especially if it contains sensitive details like political belief or voter preference. If that information is misappropriated, improperly disclosed, or otherwise exploited by political parties, there is a chance that citizens could simply disengage altogether, removing themselves from the new digital environments that could support the deepening of democratic practice.

In all fields, from social service to private business, it is well documented that constituent loyalty and engagement depends heavily on the security and confidence of the constituent-service provider relationship. A 2009 FIPA study into British Columbia's Integrated Case Management system, for example, found that those who rely on certain medical and social services are far less likely to access critical resources if there is a

perceived threat to the security of their personal relationship with a service provider. Similarly, the protection of customer privacy is seen as a critical component of any brand or corporate strategy. A 2001 report by major corporate consulting firm PricewaterhouseCoopers begins with this unequivocal statement: "E-privacy is a critical enabler for e-business. Far from being a hindrance or a cost of doing business online, e-privacy is now emerging as a key differentiator in the digital marketplace—and an absolute prerequisite to building the high levels of consumer loyalty and corporate trust that enable e-business to flourish."

Clearly, the protection of constituency privacy is essential in maintaining effective service relationships. The same principle holds for citizen engagement with political parties. Without statutes that protect the privacy rights of citizens and hold parties to account for their use, collection, and disclosure of voter information, we risk driving citizens away from an increasingly digital political landscape.

4. **Respect domestic privacy policy by guaranteeing 'carve outs' in international trade and cooperation agreements.**

   Developing any kind of policy is a delicate matter that demands careful balances and comprehensive research. As this submission hopefully makes clear, this is especially true when it comes to privacy policy in the age of social media. FIPA believes that it is important that these careful democratic efforts are ultimately respected and protected within Canada's broader policy climate.

   More specifically, this means that domestic privacy policies should not be undermined by other federal agreements and commitments, such as trade partnerships that promote the unaccountable "free flow" of Canadian data to or across other jurisdictions. Initiatives such as the Trans-Pacific Partnership Agreement, for example, must be adopted in such a way that the contextual privacy rights of Canadians, as defined in Canadian law, are respected. This means establishing 'carve outs' within international agreements meant to preserve the careful balances and compromises built into democratically-enacted privacy statutes.

   Such carve outs are essential in that they demonstrate respect for due process and Canadian parliamentary democracy. More importantly, they put into practice a vision of privacy as a social value that resists the unaccountable integration of Canadians' social lives and networks into comprehensive trade agreements on narrowly defined economic grounds.

5. **Review and draw upon policy frameworks already taking steps to realize privacy as a social value.**

   In many jurisdictions, government officials and privacy advocates have already begun to develop policy frameworks that balance the real benefits of connectivity through social networking environments with the need to establish privacy as a shared social value. Some examples of these efforts include:

- California's recent passage of a bill (AB-1844) that forbids employers from demanding social media credentials (passwords, usernames, image management) from prospective employees. A complementary piece of legislation, SB-1349, lays out similar provisions with respect to the relationship between colleges, universities, and prospective students. These policies are instructive in that they acknowledge that, while people often "disclose" personal information like photos, habits and tastes in social media environments, those acts of disclosure are contextually-situated and legislated by certain norms of sociality and interaction. Delaware, Maryland, and Illinois have also adopted similar (though in many ways still inadequate) initiatives, and the *Social Networking Online Privacy Act*, introduced into American Congress in April 2012, seeks to create comparable protections on a federal level.
- Privacy Commissioner Stoddardt's recent calls to amend *PIPEDA* and other Canadian privacy laws so as to make them commensurate with the world of big data, data mining, and the infrastructure of social networking. Commissioner Stoddart's strong words on the issue of social media privacy and her office's ongoing investigations into such digital corporate giants as Facebook are instructive in that they illustrate the limitations and constraints of current Canadian privacy legislation, and highlight the need for change and regular review.
- Ontario Information and Privacy Commissioner Cavoukian's "Privacy by Design" framework is also helpful on this issue. In attempting to acknowledge and preserve the benefits of socially mediated connectivity—economic growth, innovation, creativity, community building, democratic conversation—while advocating for the creation of tech infrastructure that errs toward less collection rather than more by default, the Privacy by Design framework lays out possible policy directions suitable to the age of social media.

CONCLUSION AND WORKS CITED

The B.C. Freedom of Information and Privacy Association thanks and applauds the House Standing Committee for welcoming submissions on this important issue. Privacy, particularly in the context of a socially mediated world, calls for thorough examination, collaborative research, and creative problem solving. We are pleased to have contributed to this process and hope that this submission will help you find a Canadian privacy framework that preserves the economic and cultural benefits of connection and interactivity while deepening commitments to privacy as a shared social value.

***

Cavoukian, A. (2009). *Privacy by Design: Take the Challenge.* Retrieved from http://www.privacybydesign.ca/content/uploads/2010/03/PrivacybyDesignBook.pdf

Cavoukian, A. (1998). *Data Mining: Staking a Claim on Your Privacy*. Retrieved from http://www.ipc.on.ca/images/resources/datamine.pdf

Bigus, J. (1996). *Data Mining With Neutral Networks: Solving Business Problems from Application Development to Decision Support.* New York: McGraw-Hill

Greenwald, M. (2012, October 5). Seen and Heard: Social Media and Privacy Legislation. Posted to *Public Policy and Governance Review* at http://ppgreview.ca/2012/10/05/seen-and-heard-social-media-and-privacy-legislation/

Israel, T. (2012, March 23). Bill C-12: Safeguarding Canadians' Personal Information Act-Eroding Privacy in the Name of Privacy [Web log message]. Retrieved from http://www.slaw.ca/2012/03/23/billc12-safeguarding-privacy-by-eroding-it/

Poyntz, S. (forthcoming). Eyes Wide Open: Stranger Hospitality and the Regulation of Youth Citizenship.

Personal Information Protection and Electronic Documents Act, S.C. 2000, c. 5. Retrieved from http://laws-lois.justice.gc.ca/eng/acts/P-8.6/page-1.html#h-3

Steeves, V. (2009). Reclaiming the Social Value of Privacy. In I. Kerr, V. Steeves & C. Lucock (Eds.), *Lessons from the Identity Trail: Anonymity, Privacy and Identity in a Networked Society* (pp. 191-208). Oxford University Press.

Stanley, J. (2012). Eight Problems with "Big Data." American Civil Liberties Union Website. Retrieved November 25, 2012 from http://www.aclu.org/blog/technology-and-liberty/eight-problems-big-data

Teitel, E. (2011, November 1). The New Paparazzi. *Maclean's.* Retrieved December 10, 2012, from http://www2.macleans.ca/2011/11/01/the-new-paparazzi/