CIGI JUNIOR FELLOWS
POLICY BRIEF
NO. 8 › AUGUST 2013

KEY POINTS

- The collection, retention and use of Big Data presents a number of unique challenges that increasingly undermine privacy rights.
- Canada's *Personal Information and Electronic Documents Act* (PIPEDA) has established a legal foundation for protecting the privacy rights of individuals online; however, additional security safeguards need to be in place to protect Canadian privacy rights in the age of Big Data.
- The Office of the Privacy Commissioner (OPC) has the opportunity to play a pivotal role in protecting our privacy rights in a digital age by pressuring organizations to adapt a set of guidelines and best practices when collecting, retaining, using and securing Big Data.

BIG DATA, BIG RESPONSIBILITIES: RECOMMENDATIONS TO THE OFFICE OF THE PRIVACY COMMISSIONER ON CANADIAN PRIVACY RIGHTS IN A DIGITAL AGE

SAMANTHA BRADSHAW, KYLE HARRIS AND HYLA ZEIFMAN

INTRODUCTION

Big Data is an umbrella term that encompasses the collection, retention and use of a massive volume and variety of data about individuals. Enabled by the Internet, it is collected in the process of online searches, the creation of social media accounts, surveys, and the data mining of phone calls and text message logs. While there are many positive uses for Big Data, there are also a number of associated policy challenges.

It is important that the OPC pay attention to the new and unique privacy implications associated with Big Data. Given the increased volume of its creation, usage and storage, as well as the increased velocity with which it is used and exchanged, there are direct implications for Canadian privacy rights. While the OPC is currently engaged in several digital literacy and educational campaigns, this policy brief puts forward additional recommendations to supplement the campaigns to further protect the privacy rights of Canadians.

CIGI JUNIOR FELLOWS POLICY BRIEF SERIES

The CIGI Junior Fellows program at the Balsillie School of International Affairs provides students with mentorship opportunities from senior scholars and policy makers. The program consists of research assistantships, policy brief writing workshops, interactive learning sessions with senior experts from CIGI and publication opportunities. Working under the direction of a project leader, each junior fellow conducts research in one of CIGI's program areas. This series presents those policy briefs that met CIGI's publications standards.

The Balsillie School of International Affairs is an independent academic institution devoted to the study of international affairs and global governance. The school assembles a critical mass of extraordinary experts to understand, explain and shape the ideas that will create effective global governance. Through its graduate programs, the school cultivates an interdisciplinary learning environment that develops knowledge of international issues from the core disciplines of political science, economics, history and environmental studies. The Balsillie School was founded in 2007 by Jim Balsillie, and is a collaborative partnership among CIGI, Wilfrid Laurier University and the University of Waterloo.

PRIVACY IN A DIGITAL AGE

KEY REQUIREMENTS FOR INTERNET PRIVACY

PIPEDA is the primary piece of Canadian legislation that protects individual privacy rights online. It establishes rules for managing personal information and aims to strike a balance between the right to privacy and the need for governments and organizations to collect, use and disclose personal information for legitimate purposes. Within PIPEDA is the Canadian Standards Association's (CSA) Model Code, which established eight principles for fair information practices to protect an individual's privacy (CSA, 2004):

- **Notice:** Users must be informed when data is being collected, what information is being collected, how the company will use the data, how long they will retain it and with whom they will share it.
- **Choice:** Users must be given a choice whether or not their personal data is collected.
- **Access:** Users must be able to access, edit and confirm the validity of their personal information upon request.
- **Security:** Users' data must be safeguarded from unauthorized access.
- **Scope:** Only the required user information should be collected.
- **Purpose:** Data may only be used for the purpose it was collected.
- **Limitations:** User data should not be held indefinitely.
- **Accountability:** Organizations must ensure privacy policies are followed and are liable if breached.

Despite the Model Code's implementation, its principles have not always been upheld, thereby allowing actors to extract, store and utilize personal information that puts users' privacy at risk. As a result of these shortcomings, PIPEDA has failed to strike a balance between the interests of individuals, governments, corporations and organizations. It is important that privacy-sensitive information adhere to these core principles to ensure Canadian's online privacy rights are being protected and respected.

WHAT IS PRIVACY-SENSITIVE INFORMATION?

Privacy-sensitive information is composed of information that may vary from one individual to another. However, common categories of privacy-sensitive information can be categorized as follows:

- **Personally identifiable information:** user's name, address, credit card numbers and Internet Protocol addresses.
- **Lifestyle information:** race, religion, relationship status, sexual orientation, hobbies, organizations user belongs to, political affiliations, friends and family members.
- **Behavioural data:** information regarding viewing habits, the type of sites frequented and how often they visit, the amount of time spent on specific websites and the kinds of purchases made both online and offline through store loyalty programs or credit cards.
- **Unique device identifiers:** information regarding the user's location, determined by a global unique identifier connected to mobile devices.

MECHANISMS OF DATA COLLECTION

Government and non-governmental organizations (NGOs) collect and store privacy-sensitive information from users around the globe in several different ways. While some methods are overt and direct, others are less visible and can occur without an individual's knowledge or consent. An organization may directly solicit and collect online information from individuals through "Terms of Service" (ToS) agreements, web forms and surveys, or on-site paperwork, such as health questionnaires or store loyalty programs. Alternatively, and increasingly more common, an organization may track and record an individual's Internet viewing habits through the use of HTTP cookies, Flash cookies, Web bugs and global unique identifiers.

When a user accesses a website, they leave an electronic footprint behind. Cookies, and other similar technologies, capture these footprints, recording and storing information about the user, such as their Internet service provider, the type of hardware and software they use, and behavioural information like viewing habits and length of time spent on a web page.

In addition, users can be tracked through website hosts, email services and social media websites. Social media companies play a central role in harvesting massive amounts of information about an individual by tracking their online activities via cookies and monitoring their patterns of movement (House of Commons, 2013). With new techniques for cross-referencing data, an individual's entire private life can be tracked, giving way to potential privacy breaches.

TYPES OF PRIVACY BREACHES

When organizations fail to properly adhere to the principles of privacy, and privacy-sensitive information is

not properly secured, users face several risks. These risks include an increased susceptibility to crimes such as fraud, identity theft, defamation or stalking. The most common types of privacy breaches can occur in the following ways:

- loss of physical storage media;
- inadequate server security, leading to computer network exploitation;
- inadequate personal computer security, compromising an individual's information;
- deliberate or inadvertent misuse of data by corporate actors;
- deliberate or inadvertent misuse of data by governments;
- surveillance and tracking through smartphone GPS and Wi-Fi capabilities; and
- unauthorized resale of data to third parties.

All of these privacy breaches become apparent and exemplified in the age of Big Data due to the vast amount of information that is being generated, stored and analyzed.

PRIVACY THREATS AND RISKS IN BIG DATA

BEHAVIOURAL TRACKING

Internet users are often surprised when they see advertisements geared to their exact interests or recent search history. These tastes are determined by the words users search and the length of time spent on a web page. While this process, known as behavioural tracking, can be regarded as a tacit cost in exchange for the use of a particular website, it is often done covertly. This leaves many individuals unaware as to when their personal information is being collected, and how it will be used or retained.

CLOUD COMPUTING

Cloud computing involves running applications or storing data on a remote, Internet-based server, rather than on a personal computer. While cloud computing provides users with a number of benefits, such as accessing files remotely via services such as Dropbox, there are several problems with the storage of data on cloud networks. One problem is that data stored beyond one's national jurisdiction is subject to different laws and regulations, which can lead to potential infringements on privacy and security. For example, data stored in countries with vague privacy regulations may be sold without the user's knowledge, resulting in governments and businesses in foreign countries gaining access to sensitive information, impacting Canadian companies' competitive advantage and individual human rights.

RESALE OF DATA

Once data surveillance and data-gathering tools monitor, profile and record the activities of online users, the collected information is frequently sold and exchanged to third parties over the Internet. This exchange of personal data often occurs without the individual's knowledge about where the data will end up and for what purposes it will be used. For example, stores such as Walmart and Target offer customers store-branded credit cards and loyalty programs that require them to provide personal information including their name, address and social insurance numbers (Hays, 2004). When consumers use these cards, their data is collected, recorded, mapped, processed and resold to suppliers who use this

information to determine which products are selling, who is purchasing them and the geographical location of the buyer.

While most companies purchase data as a major marketing asset, governments and NGOs have also been known to buy and sell personal information. The sale and resale of this data can pose massive privacy threats for individuals involved in humanitarian work or political activism. Furthermore, when businesses buy data from companies such as Walmart, Facebook and Twitter and resell this data in tertiary markets, it becomes increasingly difficult for individuals to track, access and ask for the removal of their personal data.

ToS AGREEMENTS

These agreements often give companies permission to collect and store an individual's data. Different websites have different ToS agreements that explicitly or inexplicitly state what type of data is being collected. They pose a threat to privacy when complex language is used in order to give an organization access to a customer's data for use beyond the primary purpose of the website. In these instances, many users are unable decipher the rights they are sacrificing in order to use online services. When ToS agreements are written more clearly, users must make a trade-off between using a service and sacrificing their privacy rights.

Recognizing that many websites (such as Facebook) are not free, but are rather a means to commercialize access to personal information (House of Commons, 2013), one must ask whether or not ToS agreements should be allowed to insert and enforce provisions that allow for the resale of data, and what limitations must be set on data collection and sharing, especially when minors are involved.

THE COLLECTION OF GEODATA BY MOBILE PHONES

With the increasing number of smartphone technologies incorporating GPS and Wi-Fi capabilities, it is possible to track the locations and movements of individuals on a large scale. While predicting human movement patterns may yield great benefits — such as reducing traffic congestion, improving urban planning or preventing the spread of disease — the ability to track, model and predict human movements can have several implications for privacy and security. With the collection, retention and use of geodata, it becomes easy to know the exact time and location of an individual's whereabouts and who is around them, as well as the areas they frequent. If this data is not properly secured it can compromise private life and place an individual at a higher risk of being targeted by criminals. This is especially problematic for children, as the loss and use of geodata could make them more vulnerable to stalking and online predators.

PROTECTION OF CHILDREN'S PRIVACY ONLINE

Gathering the personal information of children is a significant concern in data collection. Many websites gather children's information by offering them the opportunity to register with a site, join an online kids' club or enter a contest. This collected information is frequently sold to marketing companies and other businesses that target their advertising towards children. Effectively, children's use of the Internet as an interactive space can undermine their privacy and anonymity.

RECOMMENDATIONS

LIMITS ON THE STORAGE OF DATA

As PIPEDA fails to clearly address how long data should be stored for, many organizations hold and sell data for indefinite periods of time. Given the breadth of data exchange, there is a strong possibility for data to get lost, stolen or misused. We recommend that the OPC set out clear guidelines to supplement PIPEDA, indicating to organizations and users when it is appropriate to delete data. Guidelines should adhere to the notion of "the right to be forgotten," which permits users to demand the destruction of data that does not legally or legitimately need to be retained (European Commission, 2012). Guidelines can also encourage organizations to set a five-year limit on the retention of data — after which there is a stronger possibility for data to be lost, stolen or simply outdated. After five years, organizations should be advised to contact individuals in order to authorize the continued retention of their personal information.

We also recommend that the OPC establish guidelines that help data management and social media companies protect personal information by placing limits on the storage of geodata. It is recommended that the storage of geodata remove timestamps associated with an individual's location and the frequency of their visits. This will be critical for maintaining a certain level of privacy when it comes to the movements of an individual. Because children occupy a more vulnerable position in society, it is recommended that the OPC help mobile phone service providers and data management and social media companies establish clear rules that disable the collection and storage of any geodata from minors. By limiting the length of time data is stored for, and by giving individuals greater control over the storage of their data, the possibility for privacy infractions decreases.

THE LOCALIZATION OF DATA CENTRES AND DATA PASSPORTS

Due to the jurisdictional privacy and security risks associated with storing data on a cloud network, the OPC should create guidelines that reinforce the benefits of storing privacy-sensitive Canadian data on local clouds. The OPC should work with Canadian tech companies to incentivize the creation of more local data storage centres; these will benefit the Canadian economy by creating new jobs and an international comparative advantage in safe data storage centres.

In addition, the OPC can recommend that the federal government collaborate internationally with other states to develop a "digital passport" system. A digital passport would protect citizens whose data is being stored on a cloud outside their own jurisdiction by marking stored data with a domestic stamp. This unique stamp would ensure that states adhere to the privacy standards of the user's home country.

ToS GUIDELINES

Principle three in PIPEDA (2000) states that the "knowledge and consent of the individual are required for the collection, use or disclosure of personal information." To uphold Canada's values in regards to privacy and personal information, we recommend that the OPC establish guidelines that help data management and social media companies develop ToS agreements that are drafted in clear and accessible language.

It is also recommended that the OPC develop policies and agreements that place limits on the data that consumers are required to hand over in a ToS agreement. The OPC can work with data management and social media companies to create rules that prohibit the collection of all data in the case of minors. Recognizing that data

management and social media companies extract value from the collection and resale of data, it is recommended that the OPC pressure them to create "pay for privacy" options. Instead of surrendering privacy rights that are written in the ToS agreement, a user may opt to pay a monthly fee that would compensate companies for the services provided by a social media or data management company, while protecting the individual's privacy.

LIMITS ON THIRD-PARTY SHARING/RESALE OF COLLECTED DATA

In accordance with PIPEDA, individuals should have the right to request access to their personal information being held by an organization. It is recommended that the OPC establish guidelines whereby social media and data management companies clearly indicate which third parties and company affiliates are buying and accessing a user's data. Organizations that collect data must make public where data is being sold so that users can easily trace where their personal information is going and retrieve it upon request.

The resale of data on tertiary markets is a major concern that is not covered under PIPEDA. This greatly inhibits an individual's ability to access their data, verify its content and ask for removal. Each time an organization resells data on tertiary markets it becomes exponentially more difficult for individuals to track the location of their personal data. It is recommended that the OPC create guidelines that prohibit the resale of data in tertiary markets in order to maintain the right to accessibility that is stated within PIPEDA.

CONCLUSION

While PIPEDA has established the foundation of privacy rights in a digital age, Big Data presents us with a number of challenges that have far-reaching consequences if not properly addressed. We urge the OPC to work in collaboration with data management and social media companies to establish guidelines that will strike a balance between privacy rights, security and commercial interests. These guidelines will help improve mechanisms for establishing accountability towards the principles found in PIPEDA. They will help to protect and promote Canadian privacy rights in a digital age.

ACKNOWLEDEGEMENTS

The authors would like to express their sincerest gratitude to Mark Raymond, whose leadership, encouragement and contribution to this project have been invaluable. We also thank Carol Bonnett and Vivian Moser for their support in the publication process.

WORKS CITED

CSA (2004). "Model Code for the Protection of Personal Information." CAN/CSA-Q830-96.

European Commission (2012). "Commission Proposes a Comprehensive Reform of Data Protection Rules to Increase Users' Control of their Data and to Cut Costs for Businesses." Press release, January 25. Available at: http://europa.eu/rapid/press-release_IP-12-46_en.htm.

Hays, Constance (2004). "What Wal-Mart Knows About Customers' Habits." The New York Times, November 14.

House of Commons (2013). "Report of the Standing Committee on Access to Information, Privacy and Ethics: Privacy and Social Media in the Age of Big Data." April 2013. Available at: www.parl.gc.ca/.

PIPEDA (2000), "Personal Information Protection and Electronic Documents Act."

ABOUT THE AUTHORS

Samantha Bradshaw is a candidate for a master's degree in global governance at the University of Waterloo, based at the Balsillie School of International Affairs (BSIA) in Waterloo, Ontario. She graduated with a B.A. (honours) in political science and legal studies from the University of Waterloo in 2012. Her research is currently examining how Big Data confers power onto private actors involved in global food security governance.

Kyle Harris is a candidate for a master's degree in global governance at the University of Waterloo, based at the BSIA in Waterloo, Ontario. He graduated with a B.A (honours) in history from the University of Waterloo in 2006 and completed his M.A. specializing in international history from the University of Waterloo in 2007. His research interests include Internet governance, civil society, human rights and nuclear non-proliferation.

Hyla Zeifman is a candidate for a master's degree in international public policy at the BSIA. She graduated with a bachelor of public affairs and policy management from Carleton University in the Arthur Kroeger College of Public Affairs and Policy Management. Her research currently examines the use of social media and information communication technologies in post-conflict reconstruction efforts.

ABOUT CIGI

The Centre for International Governance Innovation is an independent, non-partisan think tank on international governance. Led by experienced practitioners and distinguished academics, CIGI supports research, forms networks, advances policy debate and generates ideas for multilateral governance improvements. Conducting an active agenda of research, events and publications, CIGI's interdisciplinary work includes collaboration with policy, business and academic communities around the world.

CIGI's current research programs focus on four themes: the global economy; global security; the environment and energy; and global development.

CIGI was founded in 2001 by Jim Balsillie, then co-CEO of Research In Motion (BlackBerry), and collaborates with and gratefully acknowledges support from a number of strategic partners, in particular the Government of Canada and the Government of Ontario.

Le CIGI a été fondé en 2001 par Jim Balsillie, qui était alors co-chef de la direction de Research In Motion (BlackBerry). Il collabore avec de nombreux partenaires stratégiques et exprime sa reconnaissance du soutien reçu de ceux-ci, notamment de l'appui reçu du gouvernement du Canada et de celui du gouvernement de l'Ontario.

For more information, please visit www.cigionline.org.

CIGI MASTHEAD

Managing Editor, Publications	Carol Bonnett
Publications Editor	Jennifer Goyder
Publications Editor	Sonya Zikic
Assistant Publications Editor	Vivian Moser
Media Designer	Steve Cross

EXECUTIVE

President	Rohinton Medhora
Vice President of Programs	David Dewitt
Vice President of Public Affairs	Fred Kuntz
Vice President of Finance	Mark Menard

COMMUNICATIONS

Communications Specialist	Kevin Dias	kdias@cigionline.org (1 519 885 2444 x 7238)
Public Affairs Coordinator	Kelly Lorimer	klorimer@cigionline.org (1 519 885 2444 x 7265)