Office of the
Privacy Commissioner
of Canada

Metadata and Privacy

A Technical and Legal Overview

October 2014

Table of Contents

Introduction

A number of recent events in Canada and elsewhere have raised questions about whether and how certain government agencies are collecting and using metadata in the course of their activities. Metadata collection programs in the United States and Canada have recently been the subject of much media discussion. While such data may be created and used lawfully in both the public and private sector subject to appropriate legal restrictions and conditions, there appears to be an enduring debate as to what metadata is, what it can reveal and how it should be treated in the absence of an express statutory provision. We continue to see notable individuals and various organizations taking the view that metadata is to be distinguished from actual communications content, and is therefore less worthy of privacy protection.

There are already a number of sources that touch on what "metadata" is and what it can reveal. The Office of the Privacy Commissioner of Canada (OPC) has previously discussed the privacy implications of metadata. In July 2006, we issued a Fact Sheet entitled *The Risks of Metadata*[1]. As well, in May 2013, the OPC published a research report entitled *What an IP Address Can Reveal About You*[2] that highlighted how knowledge of subscriber information, such as phone numbers and IP addresses, can provide a starting point to compile a picture of an individual's online activities. Building on this past OPC work, this paper seeks to provide a technical analysis of what metadata can reveal and an overview of how the courts have interpreted metadata.

What is "Metadata"?

Simply put, metadata is data that provides information about other data. It is information that is generated as you use technology, and lets you know the who, what, where, when, and how of a variety of activities. These can range from creating a document, making a telephone call, to conducting an online chat. In the communications context, metadata provides certain details about the creation, transmission and distribution of a message. As such, metadata can, for example, include the date and time a phone call is made or the location from which an e-mail was accessed.

We generally describe metadata as information about an electronic or digital record, but the notion of metadata is undeniably broad. Given that the recent debate on the nature and value of metadata stems from the interception of metadata associated with communications, the focus of this paper is on metadata created by Internet, wireless or wireline based communications.

As we explore below, the distinction between a "communication" or "content" on one hand, and information generated by or about that communication or content, on the other, is not that clear.

[1] Available online at: https://www.priv.gc.ca/resource/fs-fi/02_05_d_30_e.asp.
[2] Available online at: https://www.priv.gc.ca/information/research-recherche/2013/ip_201305_e.asp.

Some Examples of Metadata in the Communications Context

Every time you communicate, metadata is produced. Whether you are having a face-to-face conversation with an individual, texting, chatting online or using a telephone, some information about that communication – other than the communication itself – is generated.

With respect to Internet or telephone-based communications, here are some examples of the metadata that can be generated by some common activities:

Activity	Metadata Generated
	- Phone number of caller - Phone number(s)called - Unique serial numbers of phones involved - Time of call - Duration of call - Location of each participant - Telephone calling card numbers
	- Sender's name, email and IP address - Recipient's name and email address - Server transfer information - Date, time and timezone - Unique identifier of email and related emails (Message-ID) - Content type and encoding - Mail client login records with IP address - Mail client header formats - Priority and categories - Subject of email - Status of the email - Read receipt request
	- Your name and profile biographical information including birthday, hometown, work history and interests - Your username and unique identifier - Your subscriptions - Your location - Your device - Activity date, time and time zone - Your activities, likes, check-ins and events

- Your name, location, language, profile bio information and URL
- When you created your account
- Your username and unique identifier
- Tweet's location, date, time and time zone
- Tweet's unique ID and ID of tweet replied to
- Contributor IDs
- Your followers, following and favorite count
- Your verification status
- Application sending the tweet

- Pages visited and when
- User data and possibly user login details with auto-fill features
- URLs
- Your IP address, internet service provider, device hardware details, operating system and browser version
- Cookies and cached data from websites
- Your search queries
- Results that appeared in searches
- Pages you visit from search

What Metadata Can Reveal

Depending on the context, it is sometimes difficult to set out a precise line dividing a communication from metadata. Michael Morell, a former CIA official and member of the U.S. President's Review Group on Communications Technologies, has stated that "[t]here's not a sharp distinction between metadata and content. It's more of a continuum".[3] As the American Civil Liberties Union has recognized, information about an individual's location derived from cell phone towers, the recipient or sender of an e-mail message, or Internet purchases, for example, "may not be the contents of our communications, but they can paint a profoundly detailed picture of our lives".[4]

The line between metadata and the actual content of a communication can appear illusory. The size, shape or colour of an envelope can sometimes be quite revealing as to what message it contains. For example, the color and style of the envelope may reveal if the contents are of a business or personal nature; the return address or logo on the envelope may indicate who it is from; the stamp and postage mark can reveal the date

[3] S. Ackerman, "NSA review panel casts doubt on bulk data collection claims", The Guardian, January 14, 2014, available online at: http://www.theguardian.com/world/2014/jan/14/nsa-review-panel-senate-phone-data-terrorism.
[4] "Metadata: Piecing Together a Privacy Solution", American Civil Liberties Union of California, February 2014, at p. 3, available at: https://www.aclunc.org/publications/metadata-piecing-together-privacy-solution.

it was posted and from where; handwriting as distinguished from computer generated address may suggest that the correspondence is from an individual as opposed to a sophisticated business.

To transpose this specific example in the Internet context, a URL is a delivery instruction that specifies the address of the web page an individual is requesting. In other words, it is metadata created upon trying to visit a particular website. However, it can also be content since "requesting a web page essentially means sending a message saying 'please send me back the page found at this URL'. In addition, a single URL reveals exactly which page was sought, and thus exactly what content was received".[5]

> *"In fact, mining metadata can not only expose sensitive information about the past, it can even allow an observer to predict future actions. For example, research has demonstrated that an individual's future location and activities can be predicted by looking for patterns in his friends and associates' location history. A security expert also warned that identifying phone calls from key executives at a company to or from a competitor, an attorney, or a brokerage can reveal the potential for a corporate takeover before any public announcement is made".*
>
> *Metadata: Piecing Together a Privacy Solution, a Report by the ACLU of California, February 2014, available at: https://www.aclunc.org/publications/metadata-piecing-together-privacy-solution.*

Indeed, metadata can sometimes be more revealing than content itself. In the digital age, almost every online activity leaves some sort of a personal trace.[6] Computer scientist Daniel Weitzner considers metadata "arguably more revealing [than content] because it's actually much easier to analyze the patterns in a large universe of metadata and correlate them with real-world events than it is to go through a semantic analysis of all of someone's email and all of someone's telephone calls."[7] Even terms entered into search engines can be used to identify individuals and to reveal sensitive information about them. John Battelle has coined the term "Database of Intentions" which he describes as "the aggregate results of every search ever entered, every result list ever tendered, and every path taken as a result".[8] Battelle states that "[t]his information represents, in aggregate form, a place holder for the intentions of humankind – a massive database of desires, needs, wants, and likes that can be discovered, subpoenaed, archived, tracked, and exploited to all sorts of ends. Such a beast has never before existed in the history of culture, but is almost guaranteed to grow exponentially from this day forward".[9]

---

[5] *Ibid*., at p. 4.
[6] See *Klayman v. Obama*, "Brief *Amici Curiae* of the Electronic Frontier Foundation, the American Civil Liberties Union, and the ACLU of the Nation's Capital in Support of the Appellees" (20 August 2014), USCA Case #14-5004 at pp. 12-15, available online at: https://www.eff.org/files/2014/08/20/klayman_amicus_brief.pdf.
[7] E. Nakashima, "Metadata reveals the secrets of social position, company hierarchy, terrorist cells", The Washington Post, June 15, 2013, available online at: http://www.washingtonpost.com/world/national-security/metadata-reveals-the-secrets-of-social-position-company-hierarchy-terrorist-cells/2013/06/15/5058647c-d5c1-11e2-a73e-826d299ff459_story_1.html.
[8] John Battelle, "The Database of Intentions", November 3, 2003, available online at: http://battellemedia.com/archives/2003/11/the_database_of_intentions.php.
[9] *Ibid.*

"Telephony metadata can be extremely revealing, both at the level of individual calls and, especially, in the aggregate.

Although this metadata might, on first impression, seem to be little more than "information concerning the numbers dialed," analysis of telephony metadata often reveals information that could traditionally only be obtained by examining the contents of communications. That is, metadata is often a proxy for content.

In the simplest example, certain telephone numbers are used for a single purpose, such that any contact reveals basic and often sensitive information about the caller. Examples include support hotlines for victims of domestic violence and rape. Similarly, numerous hotlines exist for people considering suicide, including specific services for first responders, veterans, and gay and lesbian teenagers. Hotlines exist for sufferers of various forms of addiction, such as alcohol, drugs, and gambling".

> *Written Testimony of Professor Edward W. Felten, United States Senate, Committee on the Judiciary, Hearing on Continued Oversight of the Foreign Intelligence Surveillance Act, October 2, 2013,*
> *available online at: http://www.cs.princeton.edu/~felten/testimony-2013-10-02.pdf.*

A report issued by the President of the United States' Review Group on Intelligence and Communications Technologies also noted that the collection of metadata over time can reveal a great amount about an individual's private life.[10] In its report, the Review Group noted that "the record of every telephone call an individual makes or receives over the course of several years can reveal an enormous amount about that individual's private life". In recommending that the United States government terminate its program of collecting and storing bulk telephony metadata as soon as practicably possible, the report notes that government access to one's phone call records can have a chilling effect on associational and expressive freedoms, and affect the relationship between an individual and the state.[11]

Finally, some have gone further and argued that non-content metadata may even be a "private communication" under the *Criminal Code* and the *National Defence Act*. Both those statutes contain the same definition of the term "private communication",[12] and both statutes provide that the lawful interception of a private communication requires that certain strict conditions be met. Given the jurisprudence of the Supreme Court of Canada which has held that Part VI of the *Criminal Code* (relating to the interception of private communications) protects not only the communication itself, but any *derivative* of that communication that would convey its substance or meaning, some have argued that metadata can in many cases meet this particular threshold.[13]

---

[10] "Liberty and Security in a Changing World", Report and Recommendations of the President's Review Group on Intelligence and Communications Technologies, December 12, 2013, available online at: http://www.whitehouse.gov/sites/default/files/docs/2013-12-12_rg_final_report.pdf.

[11] *Ibid*., at p. 117.

[12] "private communication" means "any oral communication, or any telecommunication, that is made by an originator who is in Canada or is intended by the originator to be received by a person who is in Canada and that is made under circumstances in which it is reasonable for the originator to expect that it will not be intercepted by any person other than the person intended by the originator to receive it…"

[13] Craig Forcese, "Law, Logarithms and Liberties: Legal Issues Arising From CSEC's Metadata Program" (Forthcoming), March 2014, University of Ottawa Press, at p. 14.

These views are in contrast to the position taken by some government institutions. Some government agencies do not view metadata as akin to content information or to a communication. For example, in a court filing from February 2014, the government took the position that "metadata" means "information associated with a telecommunication to identify, describe, manage or route that telecommunication or any part of it as well as the means by which it was transmitted, but excludes any information or part of information which could reveal the purport of a telecommunication, or the whole or part of its content".[14] As well, in submissions before a Parliamentary committee, the Minister of Justice has taken the view that transmission metadata is to be distinguished from content.[15]

Metadata As "Personal Information"

Both of Canada's federal personal information protection statutes define "personal information" generally as information about an identifiable individual. The *Personal Information Protection and Electronic Documents Act* (PIPEDA) defines "personal information" to mean "information about an identifiable individual, but does not include the name, title or business address or telephone number of an employee of an organization." The federal *Privacy Act* defines personal information as meaning "information about an identifiable individual that is recorded in any form", and the definition includes a list of examples of what constitutes personal information under the Act, "without restricting the generality" of the opening words of the definition.

The definition of personal information has historically been given a broad and expansive interpretation.[16] While there are clear examples of what constitutes personal information under these statutes, information that at first glance does not appear to be about any particular individual can also, when combined with other information and in certain contexts, be personal information.

For example, in *Gordon v. Canada (Health)*,[17] the Federal Court agreed that the "province" field in a database for adverse drug reactions in Canada was in that case "personal information" under the *Privacy Act*. The Federal Court held that information is about an identifiable individual if it "permits" or "leads" to the possible identification of the individual, whether on the basis of that information alone, or when the information is combined with other information from other available sources.

Some examples of personal information in the technological context include, depending on the circumstances, forms of biometric information, such as fingerprints and voiceprints, tracking information collected from a Global Positioning System ("GPS") placed in employee vehicles, and information collected through the use of radio frequency identification ("RFID") tags to track items or individuals.[18]

Our Office has also considered that seemingly innocuous information, when viewed along with other available information, can be personal information and can sometimes provide a fairly accurate picture of one's personal activities, views, opinions, and lifestyle. For example, an Internet Protocol (IP) address can be personal information if it can be associated with an identifiable individual, and can be quite revealing about an

---

[14] *Ibid.,* at p. 12.
[15] Canada, Standing Committee on Justice and Human Rights (1 May 2014) at 1150 (Hon. Peter MacKay).
[16] *(Dagg v. Canada (Minister of Finance),* [1997] 2 S.C.R., dissenting, 403 at para 68; *Canada (Information Commissioner) v. Canada (Transportation Accident Investigation and Safety Board),* 2006 FCA; *Canada (Information Commissioner) v. Canada (Commissioner of the Royal Canadian Mounted Police*), [2003] 1 S.C.R. 66, 2003 SCC 8, at para 23).
[17] *Gordon v. Canada (Health)*, 2008 FC 258.
[18] Office of the Privacy Commissioner of Canada, Interpretation Bulletin: Personal Information (Updated October 2013), available online at: https://www.priv.gc.ca/leg_c/interpretations_02_e.asp.

individual's Internet-based activities. Indeed, as the OPC's report entitled *What an IP Address Can Reveal About You* highlights, an IP address linked with basic information about a subscriber of telecommunication services can reveal a person's interests, their leanings, with whom they associate, and where they travel, among other things.

While having enough metadata can provide a lot of valuable information connected to the same individual, amassing metadata in certain contexts can sometimes also identify the specific individual associated with that data. For example, in a process known as "social network analysis" or "contact chaining" – which involves creating a graph of the human network around any specific individual – analysts can identify everyone who is one or two degrees of separation from the individual of interest. A contact chaining map can show how everyone in the network neighborhood is connected to each other. Even working only one or two "hops" from an identified suspect can result in a rapidly expanding network of contacts, some of whom may have no knowledge of the suspects. Such an example is provided in "Connecting the Dots: Tracking Two Terrorist Suspects",[19] where by monitoring the activities of two suspected individuals, including information about phone calls made, e-mails sent, and meetings held, a picture of their personal network begins to emerge. With enough information, the target could be identified, in addition to other individuals in the target's network.

On 4 August 2006, AOL released a file containing 20 million search keywords for over 650,000 users over a 3-month period, intended for research purposes. According to AOL, there was no personally identifiable information in the released data – AOL had replaced user names with a random identification number. However, that number was the same for all searches by a given individual, which meant that individuals could be matched to their account and search history and could (in some instances) be identified, sometimes by cross referencing the data with other public information (e.g., phonebook listings). AOL later stated that the data had not been properly reviewed prior to release, acknowledging that search queries themselves can sometimes include identifiable information (e.g., names, social security numbers, addresses and other things people might search on).

In one example, after sifting through the data released, the New York Times was able to identify a particular user. The individual behind AOL User No. 4417749, who was assigned this number by AOL with the purpose of protecting the searcher's anonymity, was ultimately revealed to be a 62-year-old widow from the state of Georgia through an analysis of the search queries of this supposedly anonymous AOL user.

See http://www.nytimes.com/2006/08/09/technology/09aol.html?pagewanted=all&_r=0, and also http://www.nytimes.com/imagepages/2006/08/08/business/09aol-graphic.html.

As powerful as this example is, making decisions about people based on the information they search for online could, however, also lead to inaccurate conclusions about the individual. The woman behind AOL User No. 4417749 routinely researched medical conditions for other individuals, including information related to stopping smoking. As such, relating these medical conditions to the woman identified as AOL User No. 4417749 would be inaccurate.

Knowing where an individual is, or was, at any given time can also reveal information about an identifiable individual, including information that the individual might prefer be kept hidden. For example, a photo taken with a GPS-enabled camera can reveal the exact location and time it was taken, as well as the unique ID

[19] Valdis Krebs, "Connecting the Dots: Tracking Two Identified Terrorists", orgnet.com, available online at: http://orgnet.com/tnet.html.

number of the device, which could include a smartphone. Many devices enable this collection by default, and often users are unaware of this practice. This could have some serious consequences. For example, a whistleblower, journalist or political dissident relying on the protection of anonymity in order to report malfeasance by a corporate entity, criminal, or government may find their safety compromised by this default data collection.

Mobile devices are uniquely personal. People generally carry their devices with them wherever they go and use them for all manner of activities (e-mail, messaging, phone calls, photographs and so on). These devices can, and do, transmit increasingly precise geolocation information. This is sometimes done deliberately by individuals, such as when they check in with a location-based service such as Foursquare, but it is also often without the user's knowledge or consent (e.g., when GPS-enabled smart phones broadcast their locations). Device locations can also be computed from cell phone tower data.

There have been a number of studies or experiments over the years that have demonstrated the sensitivity and/or uniqueness of location data including:

a) *In 2010, the website "I Can Stalk U" (http://icanstalku.com, although the site is now closed) was created. It analyzed photos posted online for geo-tags (location metadata) and then displayed the specific location in addition to the tweeted message;*

b) *In 2012, anti-virus programmer John McAfee was arrested in Guatemala after a geotagged photo revealed his location; and*

c) *In 2013, researchers published the study "Unique in the Crowd: The privacy bounds of human mobility" (http://www.nature.com/srep/2013/130325/srep01376/pdf/srep01376.pdf), which showed that as little as four randomly chosen spatio-temporal points could uniquely identify an individual.*

One researcher has concluded that "[t]he way we move is so unique that four points [of location information] are enough to identify 95% of people"[20] while another has stated that "[a]ny dataset that has enough information on people to be interesting to researchers also has enough information to be de-anonymized."[21] The potential for identifying an individual from metadata increases where multiple types of metadata are combined, and are associated with other available information. Private sector firms are developing search engines that mine multiple data feeds, including social media networks and open data sources from governments and the private sector. In one experiment, a reporter, starting from a single geo-tagged tweet of a randomly selected individual, was able to determine where that individual attended school, that the individual attended French immersion school, where she hung out with friends, where she babysat, that the individual played soccer, and that the individual was probably a skier or a snowboarder.[22]

---

[20] Jason Palmer, "Mobile Location Data 'Present Anonymity Risk'", BBC.com, Mar. 25, 2013, available online at: http://www.bbc.co.uk/news/science-environment-21923360.

[21] Pete Warden, "Why You Can't Really Anonymize Your Data", O'Reilly Strata, May 17, 2011, available online at: http://strata.oreilly.com/2011/05/anonymize-data-limits.html.

[22] Gillian Shaw, "The withering of secrecy: Technology reveals your life on social media", Vancouver Sun, 31 March 2014, available online at: http://www.vancouversun.com/technology/personal-tech/secrecy+Technology+reveals+your+life+social/9676829/story.html.

The Metaphone Study

Using a small data sample (based on input from 546 volunteers over a period of only a few months), researchers at Stanford University were able to demonstrate, unambiguously, that the analysis of phone metadata can reveal highly sensitive information about individuals. Using publicly available sources to identify their contacts, and based on single calls, the researchers were able to determine that individuals were contacting health, financial and legal services as well as religious organizations, among others. As might be expected, calling patterns are even more revealing. For example, the researchers were able to infer sensitive medical conditions (e.g., multiple sclerosis, cardiac arrhythmia) and firearms ownership and, in some cases, were able to corroborate these inferences using public information sources. For more information on the study, and its results, see http://webpolicy.org and the posts concerning the Metaphone project.

Accordingly, the revelatory nature of metadata is increasingly bringing into question the view that such information is less worthy of privacy protection because it is to be distinguished from content information, and that such information is less sensitive as a result. It also brings into question the view that metadata is less worthy of privacy protection because it may already be publicly available to others in some form or another. The case law interpreting the term "metadata" supports the view that there is more to metadata than meets the eye.

One Week is Enough

In one recent experiment, researchers were able to piece together an incredibly accurate portrait of an individual's life from just one week of metadata tracked from a mobile application. After one week, the researchers were able to attach a timestamp to 15,000 records and were able to determine not only the individual's work habits and personal interests, but were also able to infer a social network based on his phone and e-mail traffic. They were able to see the websites he visited, the searches he made, and the subject, sender and recipient of every one of his e-mails. They were even able to crack his password for his Twitter, Google and Amazon accounts, making it possible to change account settings or even order items using his Amazon account (something the researchers didn't actually do). For more information on the experiment, and its results, see http://www.statewatch.org/news/2014/jul/bits-of-freedom-on-the-metadata-of-your-phone.pdf.

How Do the Courts View "Metadata"?

A review of the case law in Canada interpreting the term "metadata" reveals a relative paucity of judicial interpretation as to what "metadata" means generally, and even less so in the communications context.

Of these few cases, most of them involve civil proceedings where a party has made a request for the production of metadata embedded in a certain medium, including a computer hard drive, photograph, telephone record, or electronic document. These cases primarily deal with the extent to which metadata attached to electronic documents or records should be produced pursuant to a party's obligation to produce all relevant documents to the litigation, whether such electronic metadata can be considered "electronic

information" pursuant to the applicable rules of civil procedure, and whether such metadata is relevant and probative to the litigation at issue.[23]

However, there are some cases that arise in other contexts, including applications for orders regarding the public disclosure or preservation of documents or information,[24] application for warrants for access to certain documentation or information by the Crown, and constitutional challenges to searches and seizures of information.[25]

Courts have included as metadata a time/date stamp affixed to a letter or the "fax header", and some have described metadata as being akin to a pay stub. As one court noted: "pay stubs may contain information about the number of days worked, specific days works and hourly rate similar to metadata that contains, for example, information about dates and times when an e-mail was created and sent, or dates and times when a website was accessed."[26]

In other cases, courts have gone further in their analyses and have interpreted the term to include:

- information about telephone numbers that sent and received text messages and the times the calls and messages were made;[27]
- information associated with a telecommunication which identifies, describes, manages or routes that telecommunication or any part of that telecommunication;[28]
- information that can provide corroborating information about a document including information that someone has tried to delete or obscure;[29]
- information such as the time a file was created, which user was logged on to the system when a file was created, and how long a file was open;[30] and
- information of an inferred purpose of a document and the circumstances surrounding its creation.[31]

While the treatment of the term "metadata" in Canadian jurisprudence appears to depend on the underlying context, there seems to be a general consensus that metadata is viewed as data that provides information about other data, and, in many cases, may permit the drawing of inferences about an individual's conduct or activities. Some courts have engaged in a more comprehensive discussion of what metadata means and why access to it has become an increasingly contentious issue. For example, in *Abougoush v. Sauve*,[32] the British

---

[23] See, for example: *Peter Laushway v. Albert Messervey and Sobeys Group Inc.,* 2014 NSCA 7, aff'd 2013 NSSC 47; *Frangione v. Vandongen*, [2010] O.J. No 2337; *Warman v. National Post Company*, 2010 ONSC 3670; *Bishop (Litigation Guardian of) v. Minichiello,* 2009 BCSC 358; *Hummingbird v. Mustafa*, 2007 CanLII 39610; *Spar Aerospace Ltd. v. Aerowerks Engineering Inc.,* 2007 CarswellAlta 1156, aff'd 2008 ABCA 47; *Desagagne v. Yuen et al.,* 2006 BCSC 955; *Baldwin Janzen Insurance Services (2004) Ltd. v. Janzen*, [2006] BCJ No. 753; *Ireland v. Low*, 2006 BCSC 393; *Nicolardi v. Daley*, [2002] O.J. No. 595; *Reichman v. Toronto Life Publishing Co.,* [1988] O.J. No. 1727.

[24] *United States of America v. Fraser,* 2014 BCSC 227; With respect to American case law, see also *O'Neill v. City of Shoreline,* 2010 Wash. LEXIS 870 (Wash. Oct. 7, 2010); *Armstrong v. Executive Office of the President, Office of Admin.,* 303 U.S. App. D.C. 107, 1 F.3d 1274, 1993 U.S. App. LEXIS 20527 (Aug. 13, 1993).

[25] *R. v. Vu*, 2013 SCC 60.

[26] *Frangione v. Vandongen*, [2010] O.J. No. 2337.

[27] *Canada (Attorney General) v. B. (A.),* 2014 NLCA 8.

[28] *United States of America v. Fraser*, 2014 BCSC 227.

[29] *Warman v. National Post Company*, 2010 ONSC 3670.

[30] *Desagagne v. Yuen et al*., 2006 BCSC 955.

[31] *Big Pond Communications 2000 Inc., v. Kennedy*, [2004] OJ No. 820.

[32] *Abougoush v. Sauve*, 2011 BCSC 885.

Columbia Supreme Court recognized that intimate details of an individual's lifestyle may sometimes be revealed through the collection of metadata, even metadata collected from a digital camera:

> …the camera user's tolerance for physical activity from day to day or over several days may be inferred. More particularly, the metadata may be relevant to the plaintiff's ability to, for example, be active throughout a given day and then go walking on the beach in the evening, or it may be relevant to the plaintiff's ability to spend an evening at a nightclub until some given hour, and then tolerate swimming the next morning.

But it is not just Canadian courts that have recognized that metadata can be quite revealing. In the U.S., there are examples of cases that provide some commentary on the telling nature of metadata.

The Supreme Court of the United States itself has recently recognized how much metadata about the location of an individual can reveal about that individual. In *United States. v. Jones*,[33] the Supreme Court of the United States held that installing a Global Positioning System (GPS) tracking device on a vehicle to monitor the vehicle's movements constituted a search protected by the Fourth Amendment of the United States Constitution. In her concurring opinion, Justice Sotomayor noted the scope of information that could be gleaned from a vehicle's GPS location data: "GPS monitoring generates a precise, comprehensive record of a person's public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations". Justice Sotomayor quoted from a New York State Court decision[34] reflecting what can be gleaned from GPS data:

> Disclosed in [GPS] data… will be trips the indisputably private nature of which takes little imagination to conjure: trips to the psychiatrist, the plastic surgeon, the abortion clinic, the AIDS treatment center, the strip club, the criminal defense attorney, the by-the-hour motel, the union meeting, the mosque, synagogue or church, the gay bar and on and on.[35]

Additionally, formerly classified court decisions in the United States regarding NSA surveillance programs provide insight into the rationale of the United States Foreign Intelligence Surveillance Court (FISC) in allowing government metadata surveillance, as well as some commentary on what metadata can reveal.[36] For example, in a heavily redacted FISC decision likely rendered in July 2004, a judge upholding the constitutionality of a bulk Internet metadata collection program nonetheless recognized that bulk metadata collection imposes a "much broader type of collection than other pen register/trap and trace applications".[37] In another redacted decision, a FISC judge viewed metadata in the context of communication as "information 'about the communication, not the actual communication itself', including 'numbers dialed, the length of a call, internet protocol addresses, e-mail addresses, and similar information concerning the delivery of the communication rather than the message between two parties.'"[38]

---

[33] *United States v. Jones*, 132 S.Ct. 945 (2012).
[34] *People v. Weaver*, 12 N. Y. 3d 433, (2009).
[35] Ibid, at p. 441-442.
[36] Bryce Clayton Newell, "The Massive Metadata Machine: Liberty, Power, and ~~Secret~~ Mass Surveillance in the U.S. and Europe" (2014) Vol. 10:2, I/S: A Journal of Law and Policy for the Information Society, at pp. 490-492.
[37] [case name redacted], No. PR/TT [redacted] (FISA Ct.) at p. 80, available online at: http://www.dni.gov/files/documents/1118/CLEANEDPRTT%201.pdf.
[38] 37 [case name redacted], No. PR/TT [redacted], (FISA Ct.) at p. 1, available at http://www.dni.gov/files/documents/1118/CLEANEDPRTT%202.pdf.

As well, in *Klayman v. Obama*,[39] where the constitutionality of the U.S. government's collection of bulk telephony metadata was in dispute, the United States District Court for the District of Columbia went so far to state that metadata collection, "reflects a wealth of detail about [an individual's] familial, political, professional, religious and sexual associations." The court stated:

> What metadata is has not changed over time...but the ubiquity of phones has dramatically altered the quantity of information that is now available and, more importantly, what that information can tell the Government about people's lives. (Cell phone and text message communications are so pervasive that some persons may consider them to be essential means or necessary instruments for self-expression, even self-identification.)

These cases also show that metadata includes information that can reveal quite a bit of information about individuals despite not comprising a communication itself.

Metadata and Reasonable Expectations of Privacy

The jurisprudence of the Supreme Court of Canada suggests that individuals will enjoy a reasonable expectation of privacy in information generated from computer and Internet usage that reveals core biographical information.

In *R. v. Vu*, the Supreme Court of Canada acknowledged the extent to which personal information can be gleaned from metadata. In that case, the Court recognized that "computers store immense amounts of information, some of which, in the case of personal computers, will touch the "biographical core of personal information", and as noted above, that computers contain information that is automatically generated from which sensitive personal details can be gleaned". But the Court also stressed that computers can retain files and data even after users think that they have destroyed them, and when connected to the Internet, "computers serve as portals to an almost infinite amount of information that is shared between different users and is stored almost anywhere in the world. Similarly, a computer that is connected to a network will allow police to access information on other devices".

> *"Word-processing programs will often automatically generate temporary files that permit analysts to reconstruct the development of a file and access information about who created and worked on it. Similarly, most browsers used to surf the Internet are programmed to automatically retain information about the websites the user has visited in recent weeks and the search terms that were employed to access those websites. Ordinarily, this information can help a user retrace his or her cybernetic steps. In the context of a criminal investigation, however, it can also enable investigators to access intimate details about a user's interests, habits, and identity, drawing on a record that the user created unwittingly: O. S. Kerr, "Searches and Seizures in a Digital World" (2005), 119 Harv. L. Rev. 531, at pp. 542-43. This kind of information has no analogue in the physical world in which other types of receptacles are found".*
>
> *R. v. Vu,* 2013 SCC 60.

[39] *Klayman v. Obama*, Memorandum and Order dated December 16, 2013; see also *American Civil Liberties Union v. James R. Clapper*, Memorandum and Order dated December 27, 2013.

In *R. v. Spencer*, the Supreme Court of Canada held that a name and address of a subscriber linked with a particular IP address provided the police with more than a simple name and address; rather, it provided police with the "identity of an Internet subscriber which corresponded to particular Internet usage". The Court recognized that individuals can enjoy a reasonable expectation of privacy in information that links their identity to a piece of metadata, in that case, an IP address. The Court held that police obtaining the subscriber information matching an IP address from an ISP, without a warrant, constituted a search that was not authorized by law, and therefore, violated section 8 of the *Charter*.

> *"… the Court has taken a broad and functional approach to the question [of the subject matter of a search], examining the connection between the police investigative technique and the privacy interest at stake. The Court has looked at not only the nature of the precise information sought, but also at the nature of the information that it reveals".*
>
> *R. v. Spencer,* 2014 SCC 43.

The Spencer decision is a logical extension of the Supreme Court of Canada's earlier jurisprudence on privacy and computers and computer generated information. In *R. v. Morelli*, a majority of the Supreme Court found that computers used for personal purposes, regardless of where they are found or to whom they belong, "often contain our most intimate correspondence. They contain the details of our financial, medical, and personal situations. They even reveal our specific interests, likes, and propensities, recording in the browsing history and cache files the information we seek out and read, watch, or listen to on the Internet". In *R. v. Cole*, the Court added: "This is particularly the case where, as here, the computer is used to browse the Web. Internet-connected devices 'reveal our specific interests, likes, and propensities, recording in the browsing history and cache files the information we seek out and read, watch, or listen to on the Internet'. This sort of private information falls at the very heart of the 'biographical core' protected by s. 8 of the Charter.'"

These cases highlight that individuals' computer usage, particularly when linked to the Internet, can constitute "core biographical data" to which a high reasonable expectation of privacy attaches. That this jurisprudence also makes reference to search terms, URLs, browsing history, cache files and the unwitting creation of records throughout also suggests that there is a strong privacy interest associated with some types of metadata.

Conclusion

In many cases, courts have recognized that metadata can reveal much about an individual and deserves privacy protection, while recognizing that context matters.

Government institutions that collect or are considering collecting such information should not underestimate what metadata can reveal about an individual. The same goes for private-sector organizations that are requested to disclose such data to government institutions, including law enforcement agencies. Given the ubiquitous nature of metadata and the powerful inferences that can be drawn about specific individuals, government institutions and private-sector organizations will have to govern their collection and disclosure activities according to appropriate processes and standards that are commensurate with the potential level of sensitivity of metadata in any given set of circumstances.