A Primer on Metadata:
Separating Fact from Fiction

Ann Cavoukian, Ph.D.
Information and Privacy Commissioner
Ontario, Canada

July 2013
Acknowledgements

A great deal of work went into the creation of this paper. I would like to express my sincere thanks to Stephen McCammon for his insight and tireless efforts – without Stephen, this important work would not have been possible! Special thanks must also go to Michelle Chibba, David Weinkauf, and Hannah Draper for their valuable contributions.
Table of Contents

Introduction 1

Part I – Metadata: The Real Story 3
  What is Metadata? 3
  A Day in the Life... 3
  Metadata May Be More Revealing Than Content 4

Part II – Achieving Privacy and Security in Tandem: A Proactive Approach 6
  Surveillance: Much-Needed Safeguards 6
  We Can and Must Have Both Privacy and Security 8
  The Fallacy of “Nothing to Hide” 8
  Oversight and Accountability 9

Conclusion 11

Summary Table – Metadata: The Real Story 12
  Tackling the Top Three Myths Associated with Metadata 12
The purpose of this paper is to provide a primer on the much-flaunted term “metadata” – just what does it mean? Senior government officials are defending the sweeping and systemic seizure of the public's personal communications on the basis that it is “only metadata.” We are only gathering metadata which, they say, is neither sensitive nor privacy-invasive since it does not access any of the content contained in the associated communications. This paper responds to that argument, as well as considers the need for taking a proactive approach aimed at ensuring that governments achieve both security and privacy, in tandem, in an effort to ensure much-needed, overarching accountability.

On June 5th, 2013, the media published a secret U.S. Foreign Intelligence Surveillance Act (FISA) Court order requiring Verizon Communications to provide the U.S. National Security Agency (NSA) with all of its customers’ telephony metadata, for all communications between the United States and abroad, as well as those “wholly within the United States, including local telephone calls.” Further reports indicated that such classified orders are routinely sought and obtained with respect to virtually all U.S. telecoms. On this basis, it appears that the NSA is collecting and retaining most, if not all, metadata transiting the U.S. – with respect to every telephone, cell phone, or smartphone call, whether attempted or actually made. The surveillance machinery underlying this program, as well as the related PRISM Internet surveillance program, is now beginning to face much-needed public scrutiny.

On July 6th, the public learned that this program is being authorized by the FISA Court on the basis that it might be helpful to the NSA’s foreign intelligence mandate. The Court’s classified rulings amount to a “secret body of law” giving the NSA broad powers to amass vast collections of data on Americans and others. Geoffrey R. Stone, a professor of constitutional law at the University of Chicago, said he was troubled by the idea that “the FISA Court is creating a significant body of law without hearing from anyone outside the government, forgoing the adversarial system that is a staple of the American justice system.” It appears that the FISA Court has “quietly become almost a parallel Supreme Court, serving as the ultimate arbiter on surveillance issues.” This sweeping and secretive surveillance program is a far cry from a confidential but accountable warrant application process. Warrants are directed at specific suspects. The key legal issues are routinely debated and decided publicly in open court. As

---

1 See the FISA Court order available at: http://s3.documentcloud.org/documents/709012/verizon.pdf.
things stand, the degree of secrecy surrounding the program takes the appearance of surveillance without accountability.

On July 9th, Steven G. Bradbury, a key Justice Department lawyer behind the NSA’s call log program, defended it by arguing that it did not amount to surveillance. “Surveillance,” he said, “means content collection, not metadata collection.”3 According to whom? The word “surveillance” means “close watch kept over someone or something.” A program centred on the sweeping seizure of communications metadata would indeed fall under the definition of a surveillance program, capable of facilitating the indiscriminate monitoring of individuals.

Given the implications for privacy and freedom, it is critical that we all question the dated, but ever-so-prevalent either/or, zero-sum mindset to privacy vs. security. We seek to overturn the view that in order to have security, we must effectively give up our right to privacy. Instead, we believe that what is needed are measures designed to provide for both security and privacy, in an accountable and transparent manner.

Let us begin by taking apart the notion that since metadata does not access the content of our communications, it is not invasive of privacy. Really? No – absolutely not. Metadata can actually be more revealing than content.

---

What is Metadata?

Metadata is information generated by our communications devices and our communications service providers, as we use technologies like landline telephones, mobile phones, desktop computers, laptops, tablets or other computing devices. It is essentially information about other information, in this case, relating to our communications.

Pieces of metadata or traffic data are the digital crumbs that we leave behind when we use communications technologies and online services. Metadata includes information that reveals the time and duration of a communication, the particular devices, addresses, or numbers contacted, which kinds of communications services we use, and at what geolocations. And since virtually every device we use has a unique identifying number, our communications and Internet activities may be linked and traced with relative ease – ultimately back to the individuals involved. All this metadata is collected and retained by communications service providers for varying periods of time, including by telecommunications companies and Internet Service Providers, for an array business purposes. Key questions arise, however, including: who has access to all of this information, and for what purpose? The answers are especially important due to how revealing this information can be.

A Day in the Life...

As one blogger recently highlighted, the metadata created by the devices that two individuals use to communicate with each other can reveal a great deal – information about the dynamics of their relationship, where they live, where they work, and even what time they go to sleep, what time they wake up, and when they leave their home. For example, someone with access to this metadata could safely assume that two people shared a close relationship if their devices share the same location every night. The couple’s location can be revealed, for example, through the triangulation of the devices’ signal strength with respect to associated cellphone towers. Knowing that one of the devices travels every weekday to a different location at around 9:00a.m., and remains there until roughly 5:00p.m.

4 See http://youtu.be/_o2djiZOxyA.
before returning to its regular nighttime location, can also reveal not only the individual's place of
work, but the route taken and the mode of transportation used.

But this is just the beginning. For example, more metadata is generated by the fact that, while
at work one day, one of the individuals receives a telephone call from a primary school. Right
afterwards, the person leaves work and travels to that school. This suggests that the couple has at least
one child between the ages of four and 13. The following day, one of them makes a lengthy telephone
call to a particular type of doctor's office, and a week later, visits that same office. From this digital
tail, the agency with the metadata can assume that the person requires medical attention of the kind
practised at that specialized office. Other inferences may also be drawn, based on access to additional
metadata associated with, for example, visits to social or commercial establishments, as well as from
digital interactions with family members, colleagues, and companions. In this context, we might all
reasonably ask, what is more intimate: listening in as two people share the details of their day over
the phone, or being able to chart the activities of that day digitally, in great depth, and in considerable
detail?

Metadata May Be More Revealing Than Content

Not surprisingly, news of the existence and scale of the NSA's metadata surveillance program has led
to a storm of criticism. At the same time, the program has had its high-profile defenders. On June 6th,
2013, Senator Dianne Feinstein, a senior member of the U.S. Senate Intelligence Committee, said that
“this is just metadata. There is no content involved – in other words, no content of a communication.”
The inference here is that you shouldn’t worry about the state's access to metadata because it is of little
value – only listening to your telephone calls or accessing the actual content of your communications
would be invasive of your privacy. Right? Wrong!

Metadata surveillance programs gather and analyze private sector metadata involving the activities
of the public. In so doing, they facilitate the state's power to instantaneously create a detailed digital
profile of the life of anyone swept up in such a massive data seizure.

Once this data is compiled, detailed pictures of the lives of individuals begin to emerge that may easily
be linked to places and events. The data could also reveal people's political or religious affiliations, as
well as their personal and intimate relationships. Once such detailed pictures of affected individuals
have been produced, if their identities are not already known, they can easily be determined with a few
additional data sources.

In 2012, no less than Dr. Vint Cerf (inventor of the Internet with Robert Kahn), mentioned to me in a
conversation that “traffic data can be much more revealing than the content of our communications.”
When someone of the calibre of Vint Cerf says something like that to you, it makes quite an impression.
Similarly, in June 2013, Professor Daniel Weitzner, a principal research scientist at MIT's Computer
Science and Artificial Intelligence Laboratory, referred to metadata as being “arguably more revealing
than content” because it's actually much easier to analyze the patterns in a large universe of metadata

---

wp/2013/06/06/transcript-dianne-feinstein-saxby-chambliss-explain-defend-nsa-phone-records-program/.

6 See the study done by Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen, Vincent D. Blondel, “Unique in the Crowd: The
privacy bounds of human mobility.” Scientific Reports, 2013; 3 DOI: 10.1038/srep01376.
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 ...”

A few days later, Professor Michael Geist, Canada Research Chair in Internet and E-commerce Law at the University of Ottawa, stated: “What we have now … [is] 21st century technology for surveillance purposes, with the ability to capture huge amounts of metadata and then data-mine it to capture all kinds of data that sometimes can be more revealing than the content itself.”

In short, concerns about the state's massive seizure of metadata cannot be brushed aside simply because actual content is not being accessed. On its own, sweeping metadata seizures can have enormous privacy implications. Whenever you hear someone say that metadata or traffic data is not privacy-invasive – that scooping it up is not like eavesdropping on a telephone call or reading the contents of your email – challenge that view.


8 Quoted at 4:36 of http://www.cbc.ca/ontariotoday/2013/06/17/monday-ontarios-privacy-commissioner/.
Surveillance: Much-Needed Safeguards

Now more than ever, it is crucial that we deepen our commitment to ensuring that surveillance and information-sharing regimes do not undermine the independent oversight that secures our shared rights to privacy, freedom, and security. Intrusive surveillance tools, without adequate judicial safeguards, have been referred to as “the spore of totalitarianism.”9 While this may sound extreme, it is not without substance.

Even if such disasters appear remote or hypothetical, history has taught us that injustice and tyranny are preceded by a rising tide of intrusion upon the privacy and dignity of ordinary citizens. Certainly the citizens of Germany have not “forgotten what happens when secret police or intelligence agencies disregard privacy. It is an integral part of [their] history and gives young and old alike a critical perspective on state surveillance systems.”10

In the meantime, even in free and open societies, sophisticated and readily available technologies add a whole new dimension to the state’s power to subject its citizens to surveillance. In the words of U.S. Supreme Court Justice Brennan, they make surveillance “more penetrating, more indiscriminate, more truly obnoxious to a free society. Electronic surveillance, in fact, makes the police omniscient, and police omniscience is one of the most effective tools of tyranny.”11 The time to preserve our constitutional safeguards is now, while we enjoy a strong consensus about respect for human rights and the rule of law.

People everywhere expect and deserve privacy in their online and digital communication activities. The protection of both public safety and fundamental rights requires careful attention to the implications of the relationship between law enforcement agencies and service providers.

9 Jonathan Kay, “We all have something to hide from Big Brother, we just don’t know it yet,” National Post, June 14, 2013, http://fullcomment.nationalpost.com/2013/06/14/we-all-have-something-to-hide-from-big-brother-we-just-dont-know-it-yet/.

10 Malte Spitz, “Germans Loved Obama. Now We Don’t Trust Him,” The New York Times, June 29, 2013, http://www.nytimes.com/2013/06/30/opinion/sunday/germans-loved-obama-now-we-dont-trust-him.html. And see http://www.zeit.de/datenschutz/malte-spitz-data-retention/. In this piece called “Tell-all telephone,” Zeit magazine combined six months of Malte Spitz’s telephone geolocation data that he had obtained from German telecoms giant Deutsche Telekom “with information relating to his life as a politician, such as Twitter feeds, blog entries and websites, all of which is all freely available on the internet.” The resulting dynamic Google map-based infographic provides a dramatic sense of how revealing the geolocational elements of metadata can prove to be.

The state’s power to gather information from third parties in order to identify individuals engaged in ‘activities of interest’ must be subjected to timely, exacting, and independent scrutiny. This comes in the form of the appropriate combination of prior judicial authorization and subsequent notice, reporting, and accountability requirements.12

While the in camera or non-public oversight of the NSA’s surveillance programs provided by the FISA Court cannot be dismissed out of hand (as it does provide some degree of oversight), the constraints that cloak these programs in secrecy are vaguely reminiscent of the Star Chamber.13 Instead, we need to strengthen the hard-won protections provided by independent, open courts and Parliaments, and our right to access government information to comment on matters of public interest. The suggestion that FISA and similar programs are carefully controlled under a system of prior judicial authorization risks distorting this essential privacy safeguard beyond recognition.

While we acknowledge that national security imperatives, including those related to operational secrecy, must be recognized and provided for, what is first needed is demonstrating to a judge that suspects and their associates are reasonably believed to be connected to a national security threat. Both elected representatives and the public must be given the opportunity to examine and challenge the propriety of such surveillance programs and the sufficiency of existing safeguards. This is not to say that a government trying to gather evidence about particular suspects must be required to disclose the details of their investigation. However, as Professor Daniel Solove of the George Washington University Law School recently wrote, “secrecy at the level of an individual suspect is different from keeping the very existence of massive surveillance programs secret. The public must know about the general outlines of surveillance activities in order to evaluate whether the government is achieving the appropriate balance between privacy and security. What kind of information is gathered? How is it used? How securely is it kept? What kind of oversight is there? Are [the data collection programs] even legal?”14 These basic questions must be addressed in order to hold the government accountable for its surveillance practices.

In this context, one might then ask – what about government officials and their responsibility to uphold the Constitution? The U.S. Bill of Rights was designed and enacted in large measure to “prevent misconstruction or abuse of [the state’s] powers,” including those aimed at defending the nation and the public’s safety. In a constitutional democracy, public officials are sworn to uphold and defend constitutional principles. Citizens of a constitutional democracy are guaranteed certain rights and freedoms. In the United States and Canada, these protections include the right to freedom of speech and association, as well as the right to be free from unreasonable search and seizure.

13 As indicated on Wikipedia at https://en.wikipedia.org/wiki/Star_Chamber: “The Star Chamber (Latin: Camera stellata) was an English court of law that sat at the royal Palace of Westminster from the late 15th century until 1641. Court sessions were held in secret, with no indictments, and no witnesses. Evidence was presented in writing. Over time it evolved into a political weapon, a symbol of the misuse and abuse of power by the English monarchy and courts.”
We Can and Must Have Both Privacy and Security

Whatever happens in the days ahead, how will we reconcile the need for security and operational secrecy with the need for privacy and accountability? Sadly, governments that are willing to disregard one fundamental right, such as privacy, are all too frequently prepared to sacrifice other rights rooted in respect for the individual. For this reason, we cannot allow government organizations, including security agencies, to be shielded from much-needed public scrutiny – they must be reasonably transparent in their programs, policies, and activities.

Even if the NSA’s surveillance programs or comparable programs operating elsewhere comply with current-day law, people are right to be uncomfortable with the scale of the surveillance, and the state’s insistence on near absolute secrecy. Communications metadata may easily be transformed into personally identifiable information through a process of data linkages. This may in turn reveal who we associate with, at what times, and what locations, as well as the websites we frequent, and the associated areas of inquiry. With access to such data, “you could see which individuals, families or groups were communicating with one another. You could identify any social group and determine its major actors.”

Professor Jeffrey Rosen of the George Washington University Law School and President and CEO of the National Constitution Center recently outlined the scale of the problem. As his analysis illustrates, “massive telephone and Internet surveillance programs disclosed last month are the most recent examples” of a regrettable tendency. Repeatedly, governments have “chosen technologies, policies and laws that reveal innocent information without making us demonstrably safer.” Better laws, Professor Rosen asserted, would dispense “with the ‘trust us’ mentality ... [They would] put us in a better position to detect terrorism and other serious crimes without threatening privacy. ... In other words, [we] can ... protect privacy and security at the same time.”

The Fallacy of “Nothing to Hide”

In this context, many people think that only those with “something to hide” should be concerned about their privacy – this concept is not only false, it is highly misleading. In a free and open society, the individual is the source of the state’s legitimacy. Therefore, the state must be transparent to citizens so that they can hold it to account. Conversely, citizens require privacy from the state so that they may conduct themselves freely and be in a position to enjoy their autonomy. In a free society, individuals are not compelled to divulge all of their activities to the state – there has never been such a flawed expectation of “having to tell all.” The logic of surveillance, however, argues that the reverse is true: the citizen is a source of risk and therefore must be fully transparent to the state so the state can be in control. In the words of Professor Solove:

> When privacy is compromised… it can provide the government with a tremendous amount of power over its people. It can undermine trust and chill free speech and association. It can make people vulnerable to abuse of their information and further intrusions into their lives. Even if a person is doing nothing wrong, in a free society, that person shouldn’t have to justify

---

15 Ibid note 11.

every action that government officials might view as suspicious. A key component of freedom is not having to worry about how to explain oneself all the time.\textsuperscript{17}

Precisely – we couldn’t agree more! We should not have to second-guess our law-abiding activities, just in case we might be called upon to explain them, out of context. In short, privacy is integral to our freedom. As such, it is deserving of the most rigorous safeguards, and it is imperative that we continue to demand its protection. Of course, “most people concerned about the privacy implications of government surveillance aren’t arguing for no surveillance and absolute privacy. They’d be fine giving up some privacy as long as appropriate controls, limitations, oversight and accountability mechanisms were in place.”\textsuperscript{18}

In our view, we must not relinquish our privacy, nor do we have to. We must strive to have both security\textbf{ and} privacy, in tandem. Freedom must be preserved from both terrorism and tyranny. While eternal vigilance will be required to secure our fundamental rights, including freedom from unwarranted surveillance, our right to privacy, we remain confident that we can have both public safety and personal privacy. In the meantime, even if aspects of the emerging surveillance programs can ultimately be justified, the implications for privacy and freedom are too significant to sidestep. The intrusion on privacy is self-evident, as is the fact that some of the data may be used to plan subsequent surveillance efforts or other national security operations.

Have these programs helped to prevent terrorist attacks? U.S. Senators Ron Wyden and Mark Udall, members of the U.S. Senate Select Committee on Intelligence, have indicated that American intelligence officials have “repeatedly exaggerated [the] value [of a related Internet data program] in classified statements made ... to Congress and to a secret court that oversees national security surveillance.” They further stated that they are now also “skeptical of claims about the value of the bulk phone records program.”\textsuperscript{19} Have these programs led to human rights abuses or are they subject to adequate safeguards? At present, it is very difficult to answer these questions – too much is shielded from public scrutiny. Governments must, of course, have tools to fight terrorism, but in a free and open society, governments must also adopt measures designed to protect privacy and foster accountability.

\section*{Oversight and Accountability}

Fortunately, elected officials are already pressing for greater transparency and accountability. Legislation may yet be enacted to more clearly limit and control the state’s power to access metadata, as well as provide the public with information about overly secretive national security policies and disclosure orders. In addition, following a recent meeting between President Obama and the Privacy and Civil Liberties Oversight Board, the Board indicated that reviewing the NSA’s surveillance programs is a top priority and that it will be issuing a public report about the legality and propriety of the metadata

\textsuperscript{17} Ibid note 14.
\textsuperscript{18} Ibid note 14.
program. The President has also tasked the Director of National Intelligence, James Clapper, to “consider declassifying more details about the government’s collection of U.S. phone and Internet records.”

Here in Canada, the Privacy Commissioner of Canada, Jennifer Stoddart, has indicated that she is looking into the privacy implications for Canadians, including by conferring with Robert Décary, the Commissioner who oversees the Communications Security Establishment – the Canadian counterpart to the NSA. Data Protection Commissioners in Europe and elsewhere are also ‘watching the watchers,’ and are clearly concerned by these surveillance programs. In addition, the European Parliament has asked its Committee on Civil Liberties, Justice and Home Affairs to investigate and report on whether the NSA has monitored the phone calls or emails of any EU institutions, member states, and their citizens.

Given the scope of the surveillance and the significance of the privacy violations at issue, my office will continue to speak out to educate the public as well as engage key stakeholders in Canada, the United States, and abroad. In this globally networked age, privacy knows no bounds – it is no longer simply a local issue – it transcends borders, demanding global attention. Our focus will be on proactively bringing to light the view that in fighting terrorism, free societies must adopt measures designed to provide for privacy and accountability.

Accordingly, we urge governments to adopt a proactive approach to securing the rights affected by intrusive surveillance programs. To protect privacy and liberty, any power to seize communications metadata must come with strong safeguards directly embedded into programs and technologies, that are clearly expressed in the governing legal framework. The purpose, scope, and duration of data collection must be strictly controlled. More robust judicial oversight, parliamentary or congressional controls, and systems capable of providing for effective public accountability should be brought to bear. The need for operational secrecy must not stand in the way of public accountability. Our essential need for privacy and the preservation of our freedoms are at stake.

Conclusion

Now, more than ever, we must join together to protect the essential need for privacy in all of our communications – citizens and politicians, privacy experts and security officials, consumers and business leaders – everyone must join in this exercise. It is possible that all of our metadata may be implicated. As mathematician and former Sun Microsystems engineer Susan Landau said when news of the NSA’s metadata program broke in early June, “In the world of business, a pattern of phone calls from key executives can reveal impending corporate takeovers. ... [Metadata] can also reveal sensitive political information, showing, for instance, if opposition leaders are meeting, who is involved, where they gather, and for how long.”21 Even wiretapping cannot match the level of detail drawn from comprehensive metadata collection and analysis. As Joss Wright, a researcher with the Oxford Internet Institute, said, it is “far worse than reading your diary, because you don’t write everything in your diary.”22

In 2012, U.S. Supreme Court Justice Sonia Sotomayor said that “I for one doubt that people would accept without complaint the warrantless disclosure to the Government of a list of every Web site they had visited in the last week, or month, or year.”23 Whether it’s Internet metadata or telephone metadata, now is the time to act. We must reject the view that security trumps privacy and liberty. Americans and Canadians, like so many other freedom-loving people, have given their lives for constitutional rights that say otherwise. We must band together and seek measures designed to provide for both security and privacy, in an accountable and transparent manner – our freedom and liberty may depend on it.

1) **MYTH:** Metadata is not a threat to privacy because it doesn’t access any content

**FACT:** You don’t have to access the content of one’s conversations to gain access to valuable information. Scooping up our metadata or information relating to our communications can be very revealing. Access to this data will reveal the details of our personal, political, social, financial, and working lives. It provides the raw material for the creation of detailed, comprehensive, time-stamped map-lines of who is communicating with whom, when, how often, and for how long; where the senders and recipients are located; who else is connected to whom, and so forth. Over time, our metadata can become even more revealing than the actual content of our emails or phone calls, through a process of constructing intricate social graphs.

2) **MYTH:** If you have nothing to hide, you have nothing to fear

**FACT:** Everyone should be uncomfortable about unchecked state surveillance. Why, you ask? You say, “I have nothing to hide.” Neither do I. Neither do most of us. But out of context, perfectly innocent information can take on a menacing tone. In addition, history tells us that governments willing to sacrifice the right to privacy have also been willing to sacrifice other fundamental rights rooted in respect for the individual. As such, privacy is deserving of the most rigorous safeguards – it is imperative that we continue to seek its protection.

But first things first – privacy isn’t about hiding; it’s not about secrecy. Privacy is all about control – an individual’s personal control and freedom of choice. Recall that privacy is integral to our freedom. In a free and open society, individuals must be free to make informed choices about their lives, including when and to what extent they wish to reveal the personal details of their lives. Governments, on the other hand, must be accessible and transparent to citizens – not shielded from public scrutiny. The logic of surveillance argues that the opposite is true – that the individual is dangerous and therefore must be monitored and managed by the state. Does that sound like freedom to you?

3) **MYTH:** If you want to be secure, you have to give up your privacy

**FACT:** It is critical that we reject this dated, either/or, zero-sum view. Not only is it possible to have security while protecting privacy, it is far preferable because that is the essence of freedom – going about your daily activities without fear of the state looking over your shoulder. As such, state surveillance powers must be subject to prior judicial authorization, with strong safeguards directly embedded into the programs themselves, in an accountable and transparent manner. By doing so, the state can conduct necessary surveillance, while individuals can enjoy their privacy and freedom.

**REMEMBER TO SPEAK OUT:**

“IT’S NOT JUST METADATA, IT’S MY DATA!”