OFFICE OF THE
INFORMATION & PRIVACY
COMMISSIONER
for
British Columbia

National Security Claims & Transparency Respecting Privacy Practices[1]

Privacy & Security: Disclosure
6th Annual Privacy & Security Workshop
University of Toronto

November 3 & 4, 2005

David Loukidelis
Acting Information and Privacy Commissioner for British Columbia

---

> If information technology continues unfettered, then use of the digital persona [the digital construct of an individual] will inevitably result in impacts on individuals which are inequitable and oppressive, and in impacts on society which are repressive…Focussed research is needed to assess the extent to which regulation will be sufficient to prevent and/or cope with these threats. If the risks are manageable, then effective lobbying of legislatures will be necessary to ensure appropriate regulatory measures and mechanisms are imposed. If the risks are not manageable, then information technologists will be left contemplating a genie and an empty bottle.[2]

INTRODUCTION

The subject of this conference is privacy, security and transparency. The thrust of this paper is to suggest that, the neutrality of technology notwithstanding, Canadian privacy laws applicable to data mining for national security require considerable re-thinking, notably as regards transparency and accountability for state privacy practices, if our rights to privacy are to remain meaningful in the face of rapid technological changes.

Review of our privacy laws is also urgent given the inevitable adoption in Canada of information technologies to exploit our personal information in the interests of national security. The technology and practice of data mining in the interests of national security serve in this paper to illustrate the challenges to privacy, and privacy laws, raised by information technology used for national security purposes.[3]

TERRORISM, FEAR & NATIONAL SECURITY

Each terrorist attack seems to prompt renewed calls for tougher laws and more surveillance. Canadians appear to agree. During the summer of 2005, elected officials told Canadians to better prepare themselves psychologically for a terrorist attack and Canada's top military commander referred to our homeland as a theatre of military operations. CSIS asserts that the "most significant threat to Canada is that posed by terrorism."[4]

CSIS's perspective is not surprising. Its mandate, after all, is to protect national security. Nonetheless, Canadian public health officials awaiting the next, inevitable, flu pandemic might beg to differ on whether terrorism is the "most significant threat" facing Canada. Concern about terrorism is nonetheless a prominent, if not the primary, driver of public policy in the areas of justice and of public safety. Without counselling indifference or laxness in responding to the threat of terrorist acts in Canada, we should nonetheless question whether Canada's public policy discourse has become too intensely focussed on terrorism. The author's review last year of the US State Department's annual reports on global terrorism for 1998 through 2003 indicated that, throughout those six years, terrorist acts killed some 6,800 people of all nationalities worldwide, with almost half being victims of September 11.[5]

Each and every one of these crimes is senseless and barbaric. We must undoubtedly take steps to deter terrorists from committing crimes and we must find, convict and punish them when they do manage to commit crimes. As Minister of Justice Irwin Cotler and others have argued, terrorism threatens our fundamental rights to life, liberty and security of the person. Security against criminal violence (terrorist or otherwise) is part of our human rights framework.

But we must not react out of fear or pander to fear. State responses to terrorism must realistically address the scale and nature of the threats. This vital point was well expressed in early 2005 by Justice Michael Kirby, of the High Court of Australia:

> The times now are different. The risks have changed. The technology is new. The weapons are in some ways more perilous. Control over them is more disparate. But the need for prudence and care against over-reacting is as strong today….[W]e must keep in perspective the powers of those presently ranged against the Western democracies. This is not a reason for complacency over national security or indifference to violence and risks of violence. But it is a reason for keeping our feet firmly planted on the Australian ground. We should never forget that, to the extent that we exaggerate the risks to national security, we fall into the hands of those who threaten our constitutionalism. To the extent that their threats propel us into demolishing the fundamentals of our liberal democracy, we reward the enemies of our form of government with success. To the extent that we over-react, we proffer the terrorists the greatest tribute.
>
> …

> …[M]y first message is one of proportion. We should found our policies and laws on national security upon sound data alone. We should maintain our prudence, as we have in the past….[6]

The point of terrorism, of course, is to create terror. The public has little information with which to realistically assess the risks of terrorist attack at home. Their uncertainty nurtures the fear that is sown by graphic, in-your-living-room, coverage of terrorist atrocities elsewhere. It has been observed that, as a result, elected officials are likely to feel they have few options in dealing with terrorism. They can hardly be seen to be doing nothing, even though citizens are still far more likely to die in traffic accidents or in accidents in the home than in a terrorist attack. The imperative to action, and thus possible over-reaction, exists because, if terrorists strike and officials are judged to have been idle, public outrage at their apparent negligence or callousness will be extreme. By contrast, if officials are seen to have taken steps to prevent attacks and none occur, they will be credited with having successfully fought terrorism, even if the measures they took had nothing to do with the absence of attacks.[7]

This is not to suggest for a moment that officials will try to protect Canadians for political reasons. We can expect that they will act sincerely and responsibly in discharging their duty to protect citizens. Still, it is plain that the politics of 'doing nothing' are abysmally risky for our officials and therein lies an incentive for more aggressive and widespread measures that may have little to do with actual risk. It is therefore critically important that our officials—both elected and appointed—rigorously apply a rule of rational proportionality in addressing terrorist risks. It is vital that they fashion policy and legislative responses to terrorism as dispassionately, rationally and proportionately as possible.

Have Canadian legislators succeeded in doing this? Certainly, a number of the legislative measures adopted by Parliament since 9/11 have been supported both by the general public and by commentators. Yet Parliament has, since 9/11, passed or amended laws in ways that blur the lines between the collection, use and disclosure of personal information for national security purposes and its collection, use and disclosure for other purposes, including what could be called 'ordinary' law enforcement uses. These laws, passed in the name of national security, have made it easier for state agencies to collect personal information for national security purposes and then to use it for other purposes. Amendments since 9/11 empower state officials, in the name of national security, to compel businesses and other private sector organizations to turn over customer information for national security purposes and, sometimes, for secondary law enforcement uses.[8]

Against this background of changes to our laws, we have to remember that democracies depend on clear and effective rules that both reflect the essential values of a free society and that are suited to the state activities they are intended to govern. In considering the impact of national security laws on our rights and freedoms, we must remember the risks in blurring the distinctions between national security and ordinary law enforcement laws and activities. Clear distinctions are especially vital in light of the nature, history and likely future of intelligence-gathering activities and

uses, which by their very nature are clouded in secrecy and often enjoy greater leeway in the balance with our rights and freedoms. This is especially vital when one remembers the proposition that a defining characteristic of police states is the blurring of distinctions between law enforcement and national security functions, the danger being that the rule of law eventually gives way to arbitrary decision-making by the authorities, which, however well-intended it may be, can have grave consequences for citizens' rights.

We must urgently turn our attention, moreover, to the implications of information technology for individual privacy and other rights are profound. The days of privacy through practical obscurity are gone or close to it and such privacy protection as once was offered by inefficiencies in information storage and distribution will almost certainly vanish in the coming years:

> …New technologies that provide easy access to distributed data and efficiency in processing are obviously challenging to a system that is at least partially based on protecting certain rights by insisting on inefficiencies. On the one hand there is a need to "connect the dots" and on the other hand the notion of a free society is at least partially built on keeping the power to "connect the dots" out of the control any one actor, particularly the central government. Making access to data easier and more efficient (in a sense, lowering the transaction cost of data use) magnifies and enhances government power.[9]

Using the example of data mining, the following discussion suggests that steps must be taken now to modernize our approach to privacy if it is to remain viable in the face of the enhancement of government power by from developments in information technology.

THE STATE AS AN INFORMATION CONSUMER

The privacy implications of data mining can best be understood against the backdrop of private sector aggregation and mining of personal information and the near certainty that governments will increasingly be consumers of data flowing from the private sector.

Businesses throughout the world have for some years now turned to increasingly sophisticated data analysis techniques to among other things assess credit risk, market goods and services and manage relationships to their customers. Over the past decade in particular loyalty programs have sprouted up everywhere. Some are operated by businesses on their own behalf while other programs are operated by third parties. Through these programs, businesses are collecting very detailed information about consumers' lives, habits, finances, attitudes, purchasing preferences and so on. The scope and detail of the databases is often, notably in the United States, enhanced by information gleaned from public records maintained by governments. These commercial data banks are very often for sale, with large corporations such as ChoicePoint and Axiom offering those willing to pay an increasingly wide array of information products about consumers.[10]

As mentioned earlier, since September 11, there has been a trend in Canada toward enhanced state powers to compel production of personal information for national security purposes. Governments also appear to have an increasing appetite for personal information acquired from commercial personal information databases. To date, evidence that governments are becoming consumers of personal information from commercial sources is found mostly in the United States.[11] US surveillance initiatives have started to use both public and private sector information to create powerful databases that can be mined for intelligence. These initiatives include the public-private national security and law enforcement surveillance partnership known as MATRIX and the Pentagon's (now defunct) Total Information Awareness research project.[12] They also include CAPPS I,[13] CAPPS II and Secure Flight. National security programs involving personal information will undoubtedly continue to roll out in the US and, as the announcement of Transport Canada's Passenger Protect no-fly list initiative shows, will soon arrive in Canada.[14]

It would be naïve to think that Canadian national security and law enforcement agencies will long be able to resist tapping into the ever-richer trove of digital personal information that exists in the private sector. FINTRAC, for example, has been given the authority to acquire information from commercial databases. The Passenger Protect initiative will be another Canadian example of state use of private sector data for transportation and national security purposes. Passenger Protect will surely depend in large part on passenger information collected from the private sector.

The assertion that more data means better intelligence is hard to resist, however doubtful it may be as a general proposition. Yet, as commercial databases continue to proliferate, as they become more and more comprehensive and detailed, and as data storage becomes cheaper and cheaper (tending to make databases life-long in scope), it will be very difficult for the state to resist exploiting the rich lodes of data found in the private sector, never mind in the public sector.[15]

DATA MINING

Governments will want our personal data in order to use increasingly powerful computer technologies to create knowledge. Computers can be used in a variety of ways to derive knowledge from analysis of data using bespoke or off-the-shelf software. These techniques are generally referred to as 'data mining' and they are already in widespread commercial use in Canada and elsewhere.[16] The Congressional Research Service has defined data mining this way:

> Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods (algorithms that improve their performance automatically through experience, such as neural networks or decision trees). Consequently, data mining consists of more than collecting and managing data, it also includes analysis and prediction.

> Data mining can be performed on data represented in quantitative, textual, or multimedia forms. Data mining applications can use a variety of parameters to examine the data. They include association (patterns where one event is connected to another event, such as purchasing a pen and purchasing paper), sequence or path analysis (patterns where one event leads to another event, such as the birth of a child and purchasing diapers), classification (identification of new patterns, such as coincidences between duct tape purchases and plastic sheeting purchases), clustering (finding and visually documenting groups of previously unknown facts, such as geographic location and brand preferences), and forecasting (discovering patterns from which one can make reasonable predictions regarding future activities, such as the prediction that people who join an athletic club may take exercise classes).[17]

A key characteristic of data mining is that analysis of an individual's personal information can create new, secondary, information about that person. The hidden patterns and subtle relationships that data mining detects may be recorded and thus become personal information of the individual whose life is being scrutinized and analyzed. Information about an individual's credit history, credit card purchases, law enforcement record or interactions, travel habits and so on may be mined to derive evidence, or even a finding, that she or he is a possible terrorist who should be put on a terrorist watch list or be kept under surveillance. This new personal information becomes part of the swelling river of data whose channels are, in the private and public sectors, ever-changing and difficult to follow, much less control. The easier it becomes to accumulate and analyze personal information on a massive scale, the greater the potential for intentional or unintentional misuse and error. As data banks and data mining grow in sophistication and extent, each person's life will become more and more open to scrutiny, with further details becoming visible with each new advance in data analysis techniques.

On the other hand, data mining can yield benefits, for example, in the form of improved services and greater efficiency. It may be that data mining will offer benefits for national security, although there should be no assumption as to the benefits—careful research is needed in each case to establish whether benefits can be realized. Certainly, the utility of data mining for national security purposes has been recognized by the US Congress, which has recommended its use by US agencies to combat terrorism.[18]

There are, however, a variety of concerns associated with data mining, which are heightened when data mining is used by the state for national security purposes. Experts have flagged the risks for a number of years and, more recently, US government studies of data mining initiatives have also noted the risks and recommended action.[19]

These risks are generally associated with use of data mining for surveillance of individuals, groups or populations. In the case of individuals or small groups, surveillance may be predicated on suspicion derived from other sources or it may be mass surveillance. Government should assess the risks associated with data mining

for surveillance purposes—some of which are outlined below—before data mining expands in Canada, as it is likely to do, and then act to protect privacy.

DATA MINING RISKS

The privacy risks of data mining are varied in nature and significance. While there is broad consensus in the literature about what the risks are, consensus on a hierarchy of risks is not evident. For this reason, and given the overview nature of this paper, the following outline is selective—not all of the risks are mentioned, they are not presented in any particular order, and they overlap in some respects. The goal is to establish that there are risks and then recommend action, since, when these risks are realized, they can entail real and possibly serious harm to individuals.[20]

An overall concern associated with data mining—and other information technologies—is the tendency to attribute reliability or even infallibility to the products of technology. It is therefore important that the following admonition be rigorously respected when creating and operating data mining projects:

> Although these techniques are powerful, it is a mistake to view data mining and automated data analysis as complete solutions to security problems. Their strength is as tools to assist analysts and investigators. They can automate some functions that analysts would otherwise have to perform manually, they can help prioritize attention and focus an inquiry, and they can even do some early analysis and sorting of masses of data. But in the complex world of counter-terrorism, they are not likely to be useful as the only source for a conclusion or decision. When these techniques are used as more than an analytical tool, the potential for harm to individuals is far more significant.[21]

Poor data quality

The data quality problem can have a variety of causes. Missing data, fragmented data, outdated information and poorly authenticated or unauthenticated data can all contribute to error. Where data are acquired from commercial sources, data quality may suffer because the information was originally collected for purposes that do not require high accuracy.

Take an apparently trivial example from domestic life. Personal information collected through frequent shopper programs might find its way into databases exploited for national security data mining. Affinity programs do not require high assurances of identification upon enrolment and affinity cards may be shared among family, friends or mere acquaintances. Sharing of affinity cards could, for example, lead to false association of certain purchases or habits, and therefore religious beliefs, with the putative registered shopper. Moreover, the shopper's identity is likely not to have been robustly authenticated at the outset.[22] Use of such data for national security purposes, perhaps in conjunction with other flawed data, may paint an inaccurate

portrait of an individual or, as the errors multiply across the class, skew more broadly-based analyses.

Data leakage (intentional and accidental)

Like water, information flows and it will find a way to escape. Data can and often will be spilled in a variety of ways. These can include loss or theft of poorly secured servers or storage media, the hacking of systems and retention of copies of data by contractors temporarily authorized to possess the data for service-related purposes. Data leakage magnifies the risk of misuse, including through inappropriate publication of damaging information.

Data retention

The information technology phenomena that are driving development of data mining techniques also enhance the likelihood that personal information fed into and derived from data mining projects will linger for longer and longer. Data storage is becoming cheaper every day and the technologies to find and exploit archived data are advancing all the time. These factors will be partly responsible for creation of the digital personality—the digital construct of each of us that will, in important ways, mediate between our true selves and the rest of the world, notably government.[23]

What makes the description of a person in today's global data world especially worrisome is that the portrait created is not a portrait of one's true self. Our digital selves, in other words, can hardly reflect our true selves. Analysis of data can create a caricature, but it does not create a person—and the essence of privacy is maintaining your personhood. This is of more than philosophical concern. The pooling of data streams and analysis of the data can have real and costly consequences for individuals. The longer these data linger, the harder it is to correct errors or to ensure currency, particularly where the information system is a secret national security system. Even where the data are accurate, their permanent retention will raise serious problems for those who might wish, and deserve, to be able to move on with their lives. It will become more and more difficult to obscure the folly, for example, of a youthful flirtation with radical politics. Aware of the power of our digital personae, we may withdraw or tend to the anodyne. This is hardly conducive to individual fulfillment or the wellbeing of society and government.

False positives[24]

US media reported last year the Senator Ted Kennedy was told he could not board more than one domestic flight because the name T. Kennedy was on the US no-fly list, CAPPS I. His name generated a hit when run against the list, so he was banned from flying. These were false positives—he was not the T. Kennedy on the list and should not have been flagged as a security risk, even if the 'real' T. Kennedy should have been. Senator Kennedy ultimately caught his flights because he was able to persuade managers on the scene that he was not a risk. Someone not as well known might not be so fortunate. A number of examples have been reported where

individuals have been kept off flights in the US due to false positives.[25] This is more than a minor hassle for those unable to visit an ailing parent or attend a loved one's funeral. This is more than merely inconvenient for those who must fly on business. If they cannot travel when required, their jobs are in real jeopardy (and it will certainly not help an employee if the employer discovers that the employee cannot fly because she or he is on a terrorist no-fly list).

The problem of false positives is not unique to data mining, but our tendency to trust data and the scope for significant numbers of false positives promise, in combination, to make this a pressing issue. The risk of false positives is a system-design issue. If a data mining application cannot distinguish the 'noise' of ordinary behaviour from signs of possible terrorist activity, individuals will falsely be singled out for investigation or wrongly be put on watch lists. This is not to say that data mining should never be used for terrorism-related work. Rather, effective technological solutions must be found, and meaningful procedural and substantive protections must be implemented, to guard against the impact of false positives.

Function creep

It is an axiom of privacy that personal information gathered for one purpose will inevitably find other uses:

> Once the systems to access and use personal data are in place, there is an understandable interest in using those systems for other worthwhile purposes (e.g., preventing and prosecuting violent crimes). The consistent experience with data protection suggests that, over time, there is always pressure to use data collected for one purpose for other purposes. The expansive uses to which Social Security Numbers have been put are a practical example.[26]

In Canada, the two originally-intended uses for social insurance numbers have expanded to over two dozen federal government uses and the numbers are used for a myriad of other purposes in the private sector. This is not a merely trivial example, given the identifying, linking and organizing power of the social insurance number.

In the context of data mining for national security purposes, information generated for investigative purposes, or for use on a no-fly list, might be used for ordinary law enforcement purposes or to blacklist individuals.

Blacklisting

Terrorist watch-lists are being used in the US and appear to exist in one form or another in Canada. Watch lists can have legitimate, even important, functions. Use of watch lists ought, however, to be confined to a limited scope of functions such as terrorism investigation, intelligence-gathering and security clearances. A watch list could turn into a blacklist—a list used as secret evidence, or effectively as a secret finding, to make decisions that directly affect individuals who have no knowledge of the evidence or any ability to challenge it.[27] Blacklists can, of course, be officially sanctioned or illicit. In either case, they are a concern, one that is

magnified given the risks such as poor data quality that can be associated with data mining.

Wrongful misuse of data

Concern about misuse of information derived from data mining activities is by no means unique to data mining. Examples abound from other areas, both in the public sector and the private sector.[28] Embarrassing or lucrative personal information tempts intentional misuse and the products of data mining will also be tempting.

Lack of due process

Experience with national security data mining initiatives in the US suggests that authorities can be slow to recognize the need for due process and other protections. It appears, for example, that the Transportation Security Administration has been slow to devise due process protections for those who find themselves incorrectly placed on no-fly lists in the US.[29] Yet it is critically important that individuals mistakenly placed on no-fly lists or otherwise affected by errors or abuses of data mining systems be able to obtain redress through independent, fair, simple and as transparent as possible oversight processes.

DATA MINING & CANADA'S PRIVACY LAWS

Canada's privacy laws are founded on a body of internationally-accepted fair information principles that are reflected in privacy laws throughout the world and in international instruments.[30] Our privacy laws aim to give individuals a degree of control over their own personal information throughout its life cycle. They give individuals the right to be told what information is being collected about them, who is collecting it, the uses to which it will be put, to whom it might be disclosed, and for what purposes it might be disclosed. In the private sector, the rules aim to give individuals a further degree of control by enabling them to generally choose which information to give up and for what purposes. Our privacy laws also give individuals the right to have access to their own information. They require organizations to take reasonable steps to ensure that personal information they hold and use is accurate and complete.

The following discussion illustrates how many of these rules are not fully equal to the task of meaningfully protecting privacy against risks associated with data aggregation, data sharing and data mining for national security purposes. The power of these information technologies, and the risks to individuals and society, are such that new approaches to privacy protection are required to supplement existing ones.[31]

Knowledge of collection

An axiom of privacy protection is that, with limited exceptions, individuals must be given notice of collection of their personal information at the time it is collected.

As indicated earlier, data mining almost invariably depends on collection of personal information from a variety of sources and, certainly in the national security context, this means affected individuals will usually not know of the collection of their personal information.

Some observers might suggest that this could be addressed by requiring information sellers or providers to notify affected individuals of the government's collection of information. This will be of questionable efficacy even where it is feasible at the time of collection. Further, such an indirect notice requirement is unlikely to work where personal information is collected for national security purposes, since notification will be dispensed with where national security is involved.

Notice of the purposes for collection

Another important privacy principle is that individuals are to be told the purpose for which their personal information is collected. The original collector of the information will not be collecting it for a national security purpose. This principle will therefore be honoured in the breach when the information is acquired later for national security uses. Requiring the person who originally collects the information to give notice of possible later national security use is unlikely to be acceptable to United States authorities. Nor is a notice given at the time of collection that it may be disclosed where 'required or authorized by law' sufficient.

Direct collection

Privacy laws stipulate that personal information can only be collected directly from the individual the information is about. There are exceptions to this, including for ordinary law enforcement needs—police can hardly be expected to ask a suspect for personal information needed to prosecute the suspect. The same will hold true for national security activities, meaning that indirect collection for national security data mining uses will be the norm, not the exception.

Limited collection

Although the precise standards vary somewhat, Canadian privacy laws permit organizations to collect only the personal information that is necessary for, or relevant to, the purpose for which it is collected. Where an individual's personal information that is initially collected for a commercial purpose is later used for national security data mining in conjunction with other information, the limited collection principle may have little meaning and offer inadequate protection.

Individual access

An important privacy right is the right to have access to one's own personal information. This enables individuals to find out what personal information an organization has about them, how it has been used and to whom it has been disclosed. It goes almost without saying that this right is illusory in the national security context.

Accuracy & completeness

Most Canadian privacy laws require organizations to take reasonable measures to ensure that personal information they use to make a decision affecting an individual is accurate and complete. This is not a counsel of perfection, of course, but it does require positive, ongoing efforts to ensure data quality and completeness. Unlike the other traditional rules just mentioned, this duty is meaningful in the data mining context. It is necessarily imprecise and sensibly technology-neutral, but it can be particularized on an evergreen basis at a policy and operational level. A lingering concern, however, is whether meaningful independent oversight of the design of, and compliance with, this duty is available under the present privacy protection scheme.

Independent oversight

As with any rights, rights to privacy mean little unless they can be vindicated through the rule of law. Independent oversight is a central tenet of internationally accepted privacy principles. Almost all of Canada's privacy laws provide for independent review and (to varying degrees) enforcement of privacy rights through commissioners or Ombudsmen with privacy oversight duties.[32]

PRIVACY MEASURES FOR DATA MINING

As the preceding discussion shows, privacy risks associated with data mining present challenges that our existing privacy laws are in large measure ill-equipped to meet. This is not to say that our privacy laws are irrelevant in the context of data mining and other information technologies deployed for national security purposes. To be sure, the long-standing principles of limited (and proportional) collection of personal information, use of personal information for the purpose for which it was originally collected (or a very closely related purpose), information security and independent oversight remain relevant in the context of these new technologies.

While no single approach can adequately address all risks, solutions can and must be found. There is a pressing need for Canadian governments, notably the federal government, to study the available options and move quickly to implement effective and workable legal, policy and technological measures to protect privacy. Some, but not all, of the more significant measures worth considering are now outlined. Taken together, they can to some degree meet the pressing need for legislative and

policy reform that provides for a comprehensive, one-stop approach to data mining approval and regulation for national security purposes.[33]

Data mining research

Before federal government agencies engage in data mining—with the proposed Passenger Protect flight security initiative as an example—the federal government should undertake research into the effectiveness of data mining, with emphasis on technological and other tools for enhancing privacy protection. The research should also consider legal, social and ethical issues associated with data mining.

To be clear, a central focus of this research should be whether data mining for national security purposes offers meaningful benefits that are sufficiently important to override privacy and other civil rights concerns. It was acknowledged above that data mining can be useful for national security purposes, but before any data mining initiatives proceed in Canada, it is necessary to establish that any such benefits clearly and substantially outweigh the risks for privacy and other rights and liberties and that any such risks can be properly mitigated. This is not merely an exercise in assessing the constitutionality of proposals. It is a question of responsible and proportional policy making. Data mining should not be used for national security purposes in Canada unless stringent conditions are met.

Privacy impact assessments

Many Canadian jurisdictions now have statutory or government policy requirements for a privacy impact assessment ("PIA") to be completed before a proposed program, policy or law is pursued.[34] A privacy impact assessment is a process—and an ongoing one at that—that requires an organization to assess the privacy risks of proposed programs, systems or laws and, to decide, whether they should proceed and to identify and implement mitigating measures where they do proceed. A PIA process enables privacy to be designed into new systems from the outset, thus promoting efficiency as well as good privacy practice and compliance. A mandatory PIA process, ideally with sign-off by the external oversight agency, should be a mandatory feature of any data mining governance framework.

Chief Privacy Officers for national security agencies

Large corporations now commonly have a chief privacy officer ("CPO") responsible for privacy compliance and oversight within the organization. These positions are often at the senior executive level, which recognizes the importance to a corporation's brand of good privacy practices and compliance.[35] A strong case can be made that Canada's federal, provincial and territorial governments should hire or designate CPOs in a similar fashion.

At the very least, federal agencies involved in national security and anti-terrorism activities should establish well-resourced, executive-level, CPO positions with responsibility for ensuring that information technologies such as data mining are

designed and operated lawfully. These positions would not supplant, but would liaise with, external oversight agencies such as the Privacy Commissioner of Canada and the Security and Intelligence Review Committee. The US Department of Homeland Security established a CPO position over a year ago[36] and it is time such positions were created in Canada, with executive support and real internal authority.

Prior judicial authorization for data mining activities

There should be a strong rule that data mining can be performed only on anonymized data, with identification of individuals being possible only when specified quality and cogency criteria have been met and then only with prior judicial authorization. The technology exists to do this.[37] This rule would be relevant particularly in relation to data mining undertaken at a population or large group level. Where data mining is proposed in relation to specified individuals, it should be permitted only with prior judicial authorization on the basis of particularized grounds that meet constitutional standards. These recommendations are commonly encountered in the US literature and official reports.[38]

Rules-based and other technological protections

A number of technical approaches to data mining are available to enhance privacy in data mining, while more research is required to refine other techniques before they can credibly be deployed.

Rules-based processing techniques, it has been said, offer considerable promise for privacy protection in data mining. One technique would involve use of intelligent agents (or "proof-carrying code") to centrally query distributed databases by negotiating access and permitted uses on a database-by-database basis. Where data elements might move about, they could be labelled with meta-data stipulating how the element must be dealt with. This technique would allow rules specific to particular data elements to follow the data elements. A third approach involves software applications known as 'analytical filters', which are designed to filter and discard innocent noise and retain information of interest.[39]

Audit trails

Information systems in health care and commercial applications are now commonly equipped with built-in audit systems. The best of these systems automatically log access to data files and create more or less immutable audit trails. At the most basic level, they can in real time identify when unauthorized access is attempted or succeeds. More sophisticated audit applications monitor authorized access for unusual patterns and can, either automatically or with human intervention, identify both inappropriate access and use by authorized users.

These systems enable administrators (and regulators) to ensure that rules are followed. In the context of sophisticated and powerful information technology like

data mining, strong audit capabilities are of critical importance in preventing misuses of data, data spills and even function creep.

Security of data mining systems

Although a trite proposition, data mining systems must have strong security measures in order to prevent data leakages or corruption. As noted earlier, one generally-accepted privacy principle that applies to data mining in a meaningful way is the obligation to take reasonable security measures to protect personal information against unauthorized collection, use or disclosure. This is especially important in light of the risks that can be associated with data mining by the state. Data security must be a high priority in the design and operation of data mining systems.

Due process for affected individuals

As mentioned earlier, the fact that national security is involved cannot be allowed to oust due process for affected individuals. If someone is incorrectly placed on a watch list or no-fly list, or is investigated on false premises, they should have recourse to an effective process for redress. The process should, despite the national security nature of the enterprise, be as transparent as practicable in the circumstances,[40] should be inexpensive, and should be expeditious.

Ensuring effective external oversight

Last, but by no means least, some way must be found of ensuring that there is effective, independent, oversight of data mining activities. As mentioned earlier, the Privacy Commissioner of Canada has authority to investigate privacy compliance by federal government agencies. The federal *Privacy Act*, however, is sorely in need of reform and the compliance powers of the Privacy Commissioner of Canada need to be enhanced to meet the challenges of information technology and national security. The Commissioner's powers need to be modernized and strengthened hand in hand with a substantially reworked set of legislated privacy rules that apply specifically to data mining. The Commissioner's role would complement that of the courts and Parliamentary oversight (if, in fact, enhanced Parliamentary oversight comes to pass, as it should). The Commissioner's role could also complement, in the case of CSIS, the oversight role of the Security Intelligence Review Committee.

CONCLUSION

The rights and freedoms that we have come to expect will be upheld in Canada, including our privacy rights, are not absolute. Terrorism may necessitate new strategies to protect the security of all people. Although the risk of terrorist attacks on Canada is real, government must take great care not to overstep the line. Although it may be true that, the more freedom people have, the greater the potential risks, every increase in security almost inevitably curtails rights and freedoms that are at the heart of democratic societies. Rights and freedoms that we tend to take for

granted because we have always been fortunate enough to have them can be easily eroded—in good faith or otherwise—and we must ensure that our elected officials maintain life and vibrancy in them.

No one can envy the difficult task lawmakers face in trying to strike the right balance between privacy and security, but it is critically important that they ask the hard questions and come up with appropriate answers. Meaningful reforms of Canada's privacy laws—particularly the federal *Privacy Act*—are urgently required in order to address the privacy challenges raised by data mining and other information technology applications. Those reforms are needed now.

* * *

END NOTES

[1] Portions of this paper were published in an article the author wrote that appeared in the August 19, 2005 issue of *The Lawyers Weekly*, published by LexisNexis Canada Inc., and appear with permission.

[2] R. Clarke, "Computer Matching & Digital Identity" (paper delivered at Computers, Freedom & Privacy Conference, 1993) <http://www.anu.edu.au/people/Roger.Clarke/DV/CFP93.html> (hereinafter Clarke CFP Paper).

[3] For discussion of other privacy challenges presented by post-September 11 laws and policies, see *Privacy & the USA Patriot Act—Implications for British Columbia Public Sector Outsourcing* <http://www.oipc.bc.ca/sector_public/usa_patriot_act/pdfs/report/privacy-final.pdf> (Office of the Information and Privacy Commissioner for BC, October 2004).

[4] The CSIS 2003 public report, cited by the Department of Justice at: <http://www.canada.justice.gc.ca/en/ anti_terr/threats.html (accessed September 13, 2005).

[5] Of course, hundreds more have died and been injured since then, with the attacks in Bali, Beslan, Madrid, London and Sharm el-Sheik coming to mind.

[6] M. Kirby, "Proportionality, Restraint & Commonsense" (Australian Law Reform Commission, National Security Conference, March 12, 2005). As Justices Iacobucci and Arbour wrote last year in assessing the constitutionality of s. 83.28 of the Criminal Code: "The challenge for democracies in the battle against terrorism is not whether to respond, but rather how to do so….Consequently, the challenge for a democratic state's answer to terrorism calls for a balancing of what is required for an effective response to terrorism in a way that appropriately recognizes the fundamental values of the rule of law." (*Application Under s. 83.28 of the Criminal Code (Re),* [2004] S.C.J. No. 40, 2004 SCC 42, at paras. 5 and 7.) Other distinguished jurists have issued similar warnings. As Lord Steyn put it during his 2003 FA Mann lecture: "Democracies must defend themselves….But it is a recurring theme in history that in times of war, armed conflict, or perceived national danger, even liberal democracies adopt measures infringing human rights in ways that are wholly disproportionate to the crisis….Ill conceived rushed legislation is passed granting excessive powers to executive governments which compromise the rights and liberties of individuals beyond the exigencies of the situation. Often the loss of liberty is permanent.…" J. Steyn, "Guantanamo Bay: The Legal Black Hole" (27th FA Mann Lecture, British Institute of International and Comparative Law and Herbert Smith, November 25 2003) <http://www.statewatch.org/news/2003/nov/guantanamo.pdf (accessed September 13, 2005). Lord Hoffman made similar observations in his reasons in the 2004 Belmarsh case, *A (FC) and others (FC) v. Secretary of State for the Home Department*, [2004] UKHL 56, at paras. 95 & 96.

[7] These observations have been made by a number of commentators. See, for example, B. Schneier, *Beyond Fear* (Copernicus Books: New York, 2003), at 241.

[8] For example, changes to the *Customs Act* allow border officials to require airlines to disclose advance information about arriving passengers. These changes also expanded the federal government's ability to use and share that information so that it can be used, not only for national

security purposes, but for ordinary law enforcement and other purposes. The changes also allow the federal government to share personal information about Canadians with foreign governments, without the amendments restricting information-sharing to national security uses. Also see changes to the *Aeronautics Act* and the *Personal Information Protection and Electronic Documents Act*.

[9] K.A. Taipale, "Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data", 5 Col. Sci. & Tech. Law Rev. 1, at 58-59. In *Dept. of Justice v. Reporters Committee for Freedom of Press*, 489 U.S. 749 (1989), Stevens J. said (at 780) that "there is a vast difference between the public records that might be found after a diligent search of courthouse files, county archives and local police stations throughout the country and a computerized summary located in a single clearinghouse of information."

[10] A recent news article about the consumer profiling database affiliate of Tesco, the large UK-based grocery chain, is an example of this kind of business activity. See Heather Tomlinson & rob Evans, "Tesco Stocks Up on Inside Knowledge of Shoppers' Lives", *The Guardian* (online), http://www.guardian.co.uk/business/story/0,3604,1573821,00.html (accessed September 21, 2005).

[11] The already extensive, and increasing, US federal government exploitation of commercial databases is well documented. See, for example, Daniel J. Solove, *The Digital Person* (New York University Press: New York, 2004), notably at 168-175.

[12] MATRIX stands for Multi-State Anti-Terrorism Information Exchange. To moviegoers, at least, this acronym has unfortunate echoes. Soon after its existence became public, the Total Information Awareness project was re-branded the Terrorism Information Awareness project, perhaps to avoid similarly unfortunate connotations.

[13] CAPPS stands for Computer Assisted Passenger Pre-Screening.

[14] In July of 2005, Jennifer Stoddart, Privacy Commissioner of Canada, wrote to Transport Canada officials and expressed concern about a no-fly list. On August 9, 2005, joined by other Canadian privacy commissioners, including the author, she again expressed concern about the implications of Passenger Protect.

[15] It is fair to say that commercial personal information databases in the US are richer in detail and on a much larger scale than those in Canada. This stems from at least two factors. First, there are a number of US federal privacy laws, but they cover specific sectors and are relatively generous, in part due to the long tradition of US public records laws, as regards compilation of personal information by database companies.

[16] Experts in the field, and some commentators, prefer the term 'knowledge discovery', with 'data mining' referring to a specific step in data analysis. Data mining is nonetheless the popularly used term adopted here.

[17] Congressional Research Service, *Data Mining: An Overview* (Library of Congress: Washington, 2004), at 1 (citations omitted) (hereinafter CRS Report). Data mining is fairly commonly encountered in the US federal government and promises to become more common. A May 2004 US General Accounting Office study of data mining revealed that, amongst 128 federal agencies, 52 were using or planning to use data mining. There were 131 operational and 68 planned data mining initiatives. Fourteen were related to detecting terrorist activities, 15 were aimed at detecting criminal activities or patterns and 23 at detecting fraud. See *Data Mining: Federal Efforts Cover a Wide Range of Uses*, <http://www.gao.gov/new.items/d04548.pdf> (accessed September 15, 2005) (hereinafter *GAO Data Mining General Report*).

[18] Joint Inquiry Into the Intelligence Community Activities Before and After the Terrorist Attacks of September 11, 2001 (House Permanent Select Committee on Intelligence & Senate Select Committee on Intelligence, H. Rep. No. 107-792, S. Rep. No. 107-351 (2002), at 4-6.

[19] For example, the *GAO Data Mining General Report* identified a number of commonly predicted deficiencies in US federal government data mining initiatives. In an August 2005 follow-up report, the newly re-named Government Accountability Office reported that selected agencies had taken steps to protect privacy, but privacy rights were still not being appropriately protected. See *Data Mining: Agencies Have Taken Key Steps* […] (Government Accountability Office: Washington, 2005) <http://www.gao.gov/new.items/d05866.pdf (accessed September 15, 2005). A review of data mining in and for the US Department of Defense also reported a number of deficiencies and risks. See *Safeguarding Privacy in the Fight Against Terrorism* (Report of the Technology & Privacy Advisory Committee) (US Department of Defense: Washington, 2004).

<http://www.cdt.org/security/usapatriot/20040300tapac.pdf (accessed September 15, 2005) (hereinafter TAPAC Report). The Advisory Committee recommended a number of measures to address privacy risks, some of which are discussed above.

[20] For further reading on data mining and privacy risks, see the following selected publications: R. Clarke, "Information Technology & Dataveillance", 31 Commun. ACM 5 (1988) <http://www.anu.edu.au/people/Roger.Clarke/DV/CACM88.html#Dang> (accessed September 15, 2005); TAPAC Report; GAO *Data Mining General Report*; CRS Report; Clarke CFP Paper.

[21] M. De Rosa, *Data Mining and Data Analysis for Counter-Terrorism* (Center for Strategic and International Studies: Washington, 2004), at v.

[22] Although there is likely room to improve authentication for government-issued identification this is not to suggest that affinity card programs need to do a better job of authenticating identity on enrolment. (A national identity card is not, for a number of privacy-related and efficiency reasons, the answer.)

[23] See Clarke, note 2.

[24] Of course, although it is not a privacy issue, the problem of false negatives is a serious one, since a system's failure to identify as a possible terrorist someone who actually is a terrorist can have drastic consequences.

[25] TAPAC Report, at 38-39.

[26] TAPAC Report, at 39-40. See also the classic, and still relevant, study by David H. Flaherty, *Protecting Privacy in Surveillance Societies* (Chapel Hill: University of North Carolina Press, 1989).

[27] In the 1970s, the US Senate investigation into FBI abuses, known as the Church Commission, discovered widespread use by the FBI of secret personal dossiers on American citizens whose allegiance or morals were considered suspect, with these dossiers being used to deny jobs and other opportunities (notably in the 1950s, during the McCarthy days). Timothy Garton Ash, the British historian and author, discovered long after his student days in Berlin that his visits to East German student friends had earned him a secret MI5 dossier which could have denied him public sector employment (even though MI6 had tried to recruit him in his student days). T. Garton Ash, *The File* (New York: Random House, 1997).

[28] According to the TAPAC Report, at 40, "thousands" of US Internal Revenue Service employees have been disciplined for inappropriately accessing and reviewing the tax files of well-known people. In the private sector, Bank of America employees were recently reported to have sold customer information for identity theft-related purposes. See J. Leyden, "US bank staff sold customer details", *The Register* (online) <http://www.theregister.co.uk/2005/05/24/us_banks_security_flap/ (accessed September 18, 2005).

[29] In a 2005 report, the GAO concluded that deficiencies remained in the oversight and due process aspects of Secure Flight, the latest version of the US no-fly list. See *Aviation Security: Secure Flight Development & Testing Under Way* […] (Government Accountability Office: Washington, 2005). A 2004 GAO report found related defects in CAPPS II and recommended changes. See *Aviation Security: Computer Assisted Passenger Pre-screening System Faces Significant Implementation Challenges* (Government Accounting Office: Washington, 2004). It remains to be seen whether Passenger Protect will implement due process to address errors and misuse in Canada.

[30] For example, the OECD's 1980 *Guidelines on the Protection of Privacy & Transborder Flows of Personal Data* (adopted September 23, 1980), to which Canada is a signatory.

[31] It may be objected that Canada's privacy laws already contain exceptions to the following principles. That is undoubtedly true, but it is no answer. As noted above, the nature and scale of the risks to individuals demand more.

[32] One exception to this is Nova Scotia. The Review Officer under Nova Scotia's public sector freedom of information and privacy legislation has no authority to investigate and enforce the law's privacy provisions. See the *Freedom of Information and Protection of Privacy Act*, S.N. 1993, c. 5.

[33] Whether data mining is appropriate for use for general law enforcement purposes is beyond the scope of this paper, although significant questions are raised by this prospect.

[34] For example, s. 69(5) of British Columbia's *Freedom of Information and Protection of Privacy Act*, R.S.B.C. 1996, c. 165, requires provincial government ministries to perform a PIA in accordance with standing ministerial orders respecting PIAs. An example of a PIA template, prepared by the British

Columbia government (with input from the author's office), can be found through <http://www.oipc.bc.ca/sector_public/resources/pia.htm (accessed September 18, 2005).

[35] The proposition that good privacy is good for business is forcefully proved in A. Cavoukian & T. Hamilton, *The Privacy Payoff* (Toronto: McGraw-Hill Ryerson, 2002),

[36] The CPO's home page is found at <http://www.dhs.gov/dhspublic/interapp/editorial/editorial_0338.xml (accessed September 18, 2005).

[37] Taipale, note 10, at 79-80. Also see the TAPAC Report, recommendation 2.4. Also see De Rosa, note 21, at 17-18.

[38] See, for example, Taipale, note 10, TAPAC Report, Solove, note 11. In Canada, see Arthur J. Cockfield, "The State of Privacy Laws and Privacy- Encroaching Technologies after September 11: A Two-Year Report Card on the Canadian Government", (2003-2004), 1 U. of Ottawa Law & Technology Journal 325.

[39] Taipale, note 10, at 75-78.

[40] Because of the national security interests involved, it may be necessary, where the redress process could reasonably be expected to threaten harm to national security, to permit independent representatives to examine classified information relevant to disposition of the matter. These individuals would function in ways similar to *amicus curiae*.