CHAPTER 2

Open Government Data and Confidential Commercial Information: Challenging the Future of Open Data

Teresa Scassa

Abstract

In a relatively short space of time, governments at all levels in Canada have committed to open data agendas and are making a growing volume of data available in reusable formats and under open licences. And yet even as open data advocates continue to pressure governments for greater access to more data, changes brought about by big data and artificial intelligence are affecting the open data environment. Data are a valuable commodity, and governments are increasingly entering into contracts with private-sector companies for technologies that collect, process, and analyze data. These changes raise important questions about data ownership and control. Public–private partnerships for smart cities and for algorithmic decision making by governments mean that a growing volume of data collected through public-sector activities and relied upon by governments may increasingly be in private-sector hands. This chapter considers the impact of rules that exclude confidential commercial data and commercially sensitive data from disclosure as open data. It asks: What is the future of open data where data about public activities and phenomena are increasingly privately owned?

Acknowledgements
The author gratefully acknowledges the support of the Social Sciences and Humanities Research Council of Canada for its support for the Geothink project, of which this research forms a part.

The open data movement in Canada is of relatively recent vintage, although the foundation for open government data was laid by Natural Resources Canada in its efforts to make Canadian geospatial data open in the early 2000s (Lauriault & Kitchin, 2014). In 2012, Canada signed on to the international multi-stakeholder Open Government Partnership, making open data part of its broader commitment to open government. While a few municipal governments had already taken some steps toward releasing open data, it was in the second decade of the twenty-first century that Canadians saw governments at all levels embrace a significant and more or less organized commitment to make government data open and accessible through co-ordinated open government programs. There have been important commitments to open data by provincial governments, and many of Canada's major cities now have significant collections of open data available through open data portals. The federal government has built its own open data portal and developed an open data licence (based on the UK Open Government Licence), which has been adopted in slightly modified form by some provincial governments.

The upshot of all this activity is a growing volume of government data that is available as open data through a variety of portals across the country. Accompanying the releases of data has been a mixture of hackathons, open data "book clubs," and other events designed to encourage the uptake and use of open data with a view to realizing its potential. While it is difficult to know exactly who is using open data, and to what extent, Johnson (2016, p. 4) has found that important users of open government data include "engineering consultants, property developers, post-secondary and graduate students, other government agencies and academics."

Open-data programs continue to evolve with new datasets being made available on a fairly steady basis. The future of open data could therefore be characterized as one in which, on an incremental basis, new datasets are made open, new uses are found for open data, and new user communities discover this important resource.

However, this vision may be subject to disruption by rapidly evolving information technologies that are changing both the nature and volume of data that can be collected by governments through public-sector services or activities. Such changes may also significantly impact the location of the ownership of these data. In this environment, data may be collected and processed by private-sector companies on behalf of government; alternatively, they may be collected by private-sector companies and licensed to a variety of users, which may include governments. In some cases, data about phenomena or activities once under the regulatory authority of government may be collected by private-sector companies which choose not to share these data with government, or to share only selected data under their own terms (e.g., short-term rental-economy platforms) (Scassa, 2017). Because governments generally cannot release private-sector data as open data, these changes may have a profound impact on the future of open data.

Our current context is one in which government's role as a collector (and therefore as a sharer) of data is altered and diminished by changes in technology, and by the growing role of the private sector as a source of data used by government. This chapter therefore queries the future of open data in an environment where a growing volume of data in the hands of government might not be available as open data. It begins with an introduction to open data, followed by a discussion of how two particular categories of data in the hands of government are excluded from release as open data. These categories are third-party proprietary data and confidential or commercially sensitive information. The chapter next considers how these categories may affect the availability of open data in a context in which governments increasingly contract for data, data analytics, and the hardware/software to collect and process data. Recent Canadian case law suggests that there may be an emerging role for government in managing third-party proprietary data in the public interest, which is addressed in the conclusion.

1. Opening Government Data

Ideally, open government data are made available in machine-readable formats and under open licences which impose minimal restrictions on reuse (Janssen et al., 2012). Open data have transparency and accountability value (Open Data Charter, 2015). There is also value in

open data sharing between different departments or agencies of the same government or across governments (Johnson & Robinson, 2014). In addition to these uses, open-data policies are linked to goals of stimulating innovation by providing entrepreneurs with free access to important data assets (Manyika et al., 2013). In the big data and artificial intelligence (AI) context, open data may provide datasets useful to feed analytics or to drive machine learning.

While in theory any data may be made available as open data, in reality there are some important limits on what can be made open. Thus, for example, Janssen et al. (2012, p. 258) incorporate these limitations into their definition of open data as "non-privacy-restricted and non-confidential data which is produced with public money and is made available without any restrictions on its usage or distribution." Because of privacy laws, governments generally cannot release personal information as open data or as part of an open dataset. Since the statutory definition of personal information is quite broad (e.g., "information about an identifiable individual"), this can significantly limit the availability of some datasets for release (Scassa, 2010). Identifiability is assessed in terms of any other available data, and in an era of big data, reidentification risk can be high (Rocher et al., 2019; Ohm, 2010; Sweeney, 2010). Lest privacy concerns become an overwhelming barrier to open data, activists as well as governments have been working to develop guidance and strategies for opening data in ways that protect privacy rights and avoid privacy harms (Green et al., 2017; Simperl et al., 2016; Scassa & Conroy, 2016; Borgesius et al., 2015).

Governments are also obliged not to release as open data the proprietary data of third parties, as well as any confidential or commercially sensitive data. These two distinct grounds for nondisclosure are significant restraints. A government cannot release third-party proprietary data because it is not legally entitled to do so, nor may it license such data for reuse since it does not own the content. Governments are also barred from releasing confidential commercial information, since to do so would destroy the confidential character of the information, exposing governments to liability. Commercially sensitive data cannot be released because such may damage a government's relationships with those it regulates, and with those with whom it does business, by causing those entities commercial harm. These categories of data are considered below.

1.1 Proprietary Data

In order to release data under an open licence, which authorizes virtually unrestricted use of the data, the government must "own" the data in the sense of being the party legally entitled to license them. The Open Government Licence (Canada, 2017) thus specifically excludes from the terms of the licence "third party rights the Information Provider is not authorized to license." Data in the hands of governments may come from a variety of different sources. If such data are proprietary third-party data, they are excluded from release as open data.

The concept of proprietary data requires some consideration. Confidential information is considered to be a form of intellectual property, but its status as "property" was put in question by the Supreme Court of Canada in *R. v. Stewart* (1988). In any event, as noted below, open government data typically separately precludes the release of confidential commercial information, independent of proprietary issues. Proprietary data, therefore, may also include those data over which a company claims intellectual-property rights. Although copyright law does not protect facts, it will protect a compilation that amounts to an original selection or arrangement of facts (Scassa, 2018). Thus, datasets provided to government as part of regulatory processes, or shared with government under procurement contracts or other agreements, might constitute proprietary third-party data, whether or not they are also confidential.

Not every dataset is entitled to copyright protection, and thus it can be challenging to determine whether any given dataset in the hands of government is actually the "property" of a third party. Any claim to property rights in a dataset must be based on its being an original compilation of data. Originality depends on the existence of an original selection or arrangement of data; it is not a given that these elements will be present in every compilation of data (Judge & Scassa, 2010). While the threshold for protection is low, it must still be met. Further, the protection available is not for the underlying data; rather, it is only for their original selection or arrangement. The situation is made more complex by recent case law that distinguishes between "facts" (long held by copyright law to be in the public domain) and data (Scassa, 2018). Because facts themselves remain in the public domain, copyright in a compilation of facts is

violated only if a substantial part of the selection or arrangement is taken. A "whole universe" set of facts (i.e., all the facts relating to a certain activity) may not demonstrate any original selection (Judge & Scassa, 2010). Similarly, an arrangement that is made according to an external standard will not be original. On the other hand, if data are qualitatively different from facts, datasets may be more likely to be found protectable under copyright law. Nevertheless, whether there is copyright in any given dataset is an open question. For example, the Federal Court of Appeal in *Toronto Real Estate Board v. Commissioner of Competition* (2017, para. 194) expressed the opinion that the board's compilation of data gathered from real-estate listings was not sufficiently original since its compilation amounted to a "mechanical exercise." By contrast, a compilation of seismic data was found to have copyright protection by the Alberta courts (*Geophysical*, 2017).

Notwithstanding that copyright in factual compilations is contingent, it is regularly asserted by those who create and maintain datasets. A decision not to release as open data datasets over which third-party proprietary rights have been asserted could involve some kind of process in order to assess whether proprietary rights exist, and if so, what their scope might be. The reality is, however, that where a third party has indicated to a government that their dataset is proprietary, the government will be unlikely to question this assertion. Governments' own claims to rights in their data are similarly accepted by industry. In cases where the government has contracted out for the collection or use of the data at issue, it is the contract with the third-party provider that may determine which of the parties has a proprietary claim to the data.

Claims by third parties to rights in data have clear implications for open government data. The number and importance of such claims may increase with the growth of smart cities, data analytics, and AI. Data and datasets are in high demand in the development of AI technologies. Smart cities are sensor-laden and lead to the collection of vast amounts of often very rich data that can have multiple applications, including in the AI context. Where governments have contracted with the private sector to supply the technology to collect, process, and analyze smart-cities data, or where they license data from app companies such as Waze or Strava, these data may be subject to intellectual-property rights claims by the private-sector company. If such rights exist and are located with the companies, or if the

government accepts and accedes to the claims, then the data—even if they are about public infrastructure, programs, or spaces—cannot be released as open data. This would limit the ability of other developers, researchers, and governments themselves to access and use these data for a diverse range of purposes.

1.2 Confidential and Commercially Sensitive Data

Confidential information (of which trade secrets are a subset) is often treated as a kind of intellectual property, but it is different from other forms of intellectual property. Confidential information and trade secrets depend for their protection on their confidentiality; once that confidentiality is lost, so too is their protected nature. Hagen et al. (2017) argue that, unlike other areas of intellectual property, the public benefits much less directly from the protection of confidential commercial information. This is because, rather than support the publication and/or dissemination of the protected content (as is the case with patent and copyright law), the law of confidential information aids in keeping information secret. Hagen et al. (2017, p. 579) argue that the public interest in protecting confidential information lies in part in promoting fair competition and ethical behaviour, and in part in "promoting, protecting and enforcing relationships founded on trust and confidence." Where government has received confidential commercial information from companies in the course of their regulatory role, the protection of this information enhances trust in government and contributes to frank and honest compliance with regulatory disclosure requirements.

The protection of confidential information is a matter of domestic law, but it is also addressed in important trade treaties. The characteristics of confidential information are described in Article 39(2) of the *Agreement on Trade-Related Aspects of Intellectual Property Rights* (TRIPS). Information must be protected as confidential information if it:

> (a) is secret in the sense that it is not, as a body, or in the precise configuration and assembly of components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question;
> (b) has commercial value because it is secret; and
> (c) has been subject to reasonable steps under the circumstances, by the person lawfully in control of the

> information, to keep it secret. (Agreement on Trade-Related Aspects of Intellectual Property Rights, 1994, Art. 39(2))

Article 20.72 of the *Canada–United States–Mexico Agreement* contains similar criteria. The protection of confidential information thus principally depends upon the control that is maintained over the information; this is usually achieved through physical and/or technological barriers that protect the information as well as by contracts that govern any necessary information sharing.

Governments are generally under obligations to protect from disclosure any confidential commercial information they receive from third parties. These obligations are found in domestic laws regarding access to information. For example, *Canada's Access to Information Act* (ATIA, 1985) contains a series of exceptions to the requirement to disclose information that relates specifically to third-party confidential information. Section 20 provides:

> 20 (1) Subject to this section, the head of a government institution shall refuse to disclose any record requested under this Act that contains
>
> (a) trade secrets of a third party;
>
> (b) financial, commercial, scientific or technical information that is confidential information supplied to a government institution by a third party and is treated consistently in a confidential manner by the third party;
>
> (b.1) information that is supplied in confidence to a government institution by a third party for the preparation, maintenance, testing or implementation by the government institution of emergency management plans within the meaning of section 2 of the Emergency Management Act and that concerns the vulnerability of the third party's buildings or other structures, its networks or systems, including its computer or communications networks or systems, or the methods used to protect any of those buildings, structures, networks or systems.

In addition to these exceptions to the disclosure of third-party confidential commercial information, the ATIA also carves out

exceptions for another category of information that is excluded from disclosure: commercially sensitive information. A government may refuse to disclose such information where to do so would cause commercial harm to the party that supplied it. For example, paragraphs 20(1)(c) and (d) of the ATIA direct government institutions to refuse to disclose the following types of information:

> (c) information the disclosure of which could reasonably be expected to result in material financial loss or gain to, or could reasonably be expected to prejudice the competitive position of, a third party; or
> (d) information the disclosure of which could reasonably be expected to interfere with contractual or other negotiations of a third party.

Rosenblum and Maples (2009, p. 33) describe commercially sensitive information as "any information that has economic value or could cause economic harm if known." They observe that it is a widely used basis for claiming non-disclosure of information.

It should be noted that provincial access-to-information statutes also contain exceptions to disclosure for both confidential commercial information and commercially sensitive information, but the terms of these exceptions vary. Thus, for example, under section 17(1) of Ontario's *Freedom of Information and Protection of Privacy Act* (FOIPP, 1990), a government department or agency can refuse to release "a record that reveals a trade secret or scientific, technical, commercial, financial or labour relations information, supplied in confidence implicitly or explicitly" only where certain harms might arise. These harms are found to materialize where release of the information could reasonably be expected to:

> 17 (1) […]
> (a) prejudice significantly the competitive position or interfere significantly with the contractual or other negotiations of a person, group of persons, or organization;
> (b) result in similar information no longer being supplied to the institution where it is in the public interest that similar information continue to be so supplied;
> (c) result in undue loss or gain to any person, group, committee or financial institution or agency.

Disclosure is the default position unless the applicability of one of the exceptions can be asserted. The emphasis is somewhat different in the federal statute where the default position is secrecy. Under the federal ATIA, the head of the government institution "shall" refuse to disclose trade secrets and confidential information of third parties, but it *may* disclose third-party confidential information and commercially sensitive information if:

> (a) the disclosure would be in the public interest as it relates to public health, public safety or protection of the environment; and
> (b) the public interest in disclosure clearly outweighs in importance any financial loss or gain to a third party, any prejudice to the security of its structures, networks or systems, any prejudice to its competitive position or any interference with its contractual or other negotiations. (ATIA, s. 20(6)).

This public interest override does not apply to third-party trade secrets. With the exception of Ontario (2019), there is no legislation at the provincial or federal levels that expressly addresses open data. Nevertheless, the principles relating to the release of open data track closely those found in access-to-information laws. In fact, the Ontario *Digital and Data Directive* creates an exception to the principle of open by default for: "Data that is subject to statutory confidentiality requirements, (i.e., exempt from publication under the *Freedom of Information and Protection of Privacy Act*, 1990, the *Personal Health Information Protection Act*, 2004 and/or other statutes)" (Government of Ontario, 2021). It then goes on to create a further exception for "[d]ata that should not be disclosed for legal, security, confidentiality, privacy or commercial sensitivity reasons."

The federal Directive on Open Government (Treasury Board Secretariat, 2014) describes those data that are required to be released as open data. The directive's definition of open data explicitly excludes confidential data or data owned by third parties:

> All data resources of business value held by Government of Canada departments are to be open by default and released as open data unless subject to valid exceptions, such as ownership, security, privacy, and confidentiality, as determined by the

> department. The Treasury Board of Canada Secretariat will support departments in the development of their decision-making and approval processes with regard to legal and policy issues by providing a release criteria checklist and other guidance tools. (Treasury Board Secretariat, 2014, Appendix B)

Confidential information may consist of information submitted directly by third-party companies for the purposes of regulatory compliance; or, for example, in the context of procurement or other bidding or application processes. In these contexts, the access-to-information regimes can provide guidelines about what information is public and what information is considered confidential. Companies may also draw to the attention of government specific information that is to be treated as confidential.

In the case of personal information, a combination of concerns for the public interest in privacy, as well as concerns that privacy should not unduly limit the availability of open data, have led to a considerable amount of work being done, both by governments and by those outside government, to develop tools for assessing when personal information is present in datasets and what techniques can be used to anonymize the data such that it might still be released. (see, e.g., Green et al., 2017; Simperl et al., 2016; Scassa & Conroy, 2016; Borgesius, 2015)

Interestingly, relatively little comparable work has been done in relation to confidential commercial information. Where guidance exists, it tends to be in fairly general terms (e.g., Information and Privacy Commissioner of Ontario, 2015). Certainly, the identification of confidential commercial information in a dataset is a different sort of undertaking from the identification of personal information. Whether something is confidential commercial information may well depend on the specific context and circumstances under which the information was shared with government, as well as the nature of the information and the nature of the third party's business. By contrast, with personal information, the question is more objective, even if it is sometimes challenging to answer. With personal information, the issue is whether information is about an identifiable individual. This may require an assessment of whether data can, by being linked with other available data, become information about an identifiable individual. Nevertheless, as is the case with personal information, excessive concerns over releasing confidential commercial information

could lead to reluctance to release some data as open data. An overly cautious approach could work against the public interest. This is particularly the case where the allegedly confidential data are collected in the smart-cities context, for example, and are thus data about day-to-day operations and activities within the city in which the public has a clear interest.

Case law that has evolved under federal and provincial access-to-information regimes may be helpful in understanding when information supplied by third parties qualifies as confidential commercial information or commercially sensitive information for the purposes of the exceptions to release of data under open data directives. As noted above, under the ATIA, some grounds for refusal to disclose information are mandatory; others are discretionary. The head of a government institution must refuse to disclose confidential commercial information, whereas they have the discretion to disclose commercially sensitive information. The federal court has described it in these terms:

> With respect to mandatory exemptions, there is but one question to be considered; namely, the factual determination as to whether the material comes within the description of the exempting provision. If the contested information or record is found to fall within the description set out in the mandatory exemption provision, then the head of the government institution is obligated to refuse disclosure. In the case of discretionary exemptions, there are two decisions to be reviewed by the Court. First, as with mandatory exemptions, the Court must review the record to determine if the head of the government institution erred in making the factual decision as to whether or not the requested information falls within the exempting provision. If the information meets the criteria, the Court must then determine whether the head of the government institution exercised his or her discretion properly. (*Canadian Imperial Bank of Commerce v. Canada*, 2006, para. 28)

In either case, it is necessary first to determine if the information/data at issue falls within the boundaries of the exception. In the case of the mandatory exception for confidential commercial information, a finding of confidentiality is the end of the inquiry. With the discretionary exceptions, there is a second stage, which considers whether the discretion to refuse to disclose commercially sensitive

data was properly exercised. By contrast, under Ontario's legislation, each inquiry has two steps, since even with confidential commercial information there is still discretion to disclose information in the public interest. The different approaches make it clear that, in Ontario, the protection of both confidential commercial information and commercially sensitive information can be treated as discretionary.

While, in general, courts and tribunals are responsive to the need to protect confidential or commercially sensitive information, where legislation requires third parties to establish with evidence the justification for non-disclosure, courts will critically assess whether the necessary evidence has been presented to support a claim of confidentiality (*Canadian Imperial Bank of Commerce v. Canada*, 2006; *Gartner Inc. v. Ontario*, 2017). It will not be enough for a party merely to assert that the information at issue is confidential and has always been treated as such. Courts have recognized that if the onus were not placed squarely on the third party to justify their claims to confidentiality, this "would surely undermine much of the purpose of this Act which in part is to make available to the public the information upon which government action is taken or refused" (*Ottawa Football Club v. Canada*, 1989, para. 487; *Canadian Imperial Bank of Commerce v. Canada*, 2006). In Ontario, where the confidential nature of the information is not the end of the inquiry, disclosure is refused only where the third party can demonstrate not just that the information is confidential or commercially sensitive, but also that its disclosure will create a "reasonable expectation of probable harm" (*Gartner Inc. v. Ontario*, 2017, para. 13).

In the case of the mandatory non-disclosure of confidential information under the federal ATIA, it is particularly important to ensure that claims to confidential information are carefully scrutinized. In *Air Atonabee Ltd. v. Canada* (1989, p. 272), the federal court made it clear that "whether information is confidential will depend upon its content, its purposes and the circumstances in which it is compiled and communicated." The court then went on to outline the criteria that would be considered, namely

> (a) that the content of the record be such that the information it contains is not available from sources otherwise accessible by the public or that could not be obtained by observation or independent study by a member of the public acting on his own ["not publicly available"],

> (b) that the information originate and be communicated in a reasonable expectation of confidence that it will not be disclosed ["reasonable expectation of non-disclosure"], and
> (c) that the information be communicated, whether required by law or supplied gratuitously, in a relationship between government and the party supplying it that is either a fiduciary relationship or one that is not contrary to the public interest, and which relationship will be fostered for public benefit by confidential communication ["public benefit"].

The access-to-information regime therefore establishes a balancing approach that considers the nature of the information, the relationship between the government and the party that provided it, and broader issues of public interest. It is important to note that while access-to-information regimes include both a balancing formula and checks and balances in terms of judicial review of decisions made about access, there is nothing equivalent for open data processes.

Access-to-information case law also makes it clear that it is conceivable that information or datasets may be composed of a combination of confidential and non-confidential information. Where it is possible to sever the confidential information from that which is non-confidential, it may be possible to release the non-confidential portion of the data (*Canadian Imperial Bank of Commerce v. Canada*, 2006). Thus, in the case of open data as well, it might be possible to strip proprietary or confidential commercial information from larger datasets, leaving data that are still meaningful enough to be released as open data. In general, confidential commercial information does not raise the same challenges as personal information when it comes to preparing datasets for release, since either the regulatory regime under which the information was shared or the party that shared the information will have identified that which is considered confidential. However, it is still possible for companies to be over-inclusive when it comes to identifying information as confidential commercial information. For example, in *Gartner Inc. v. Ontario* (2017), the applicant argued that the data, used in providing its benchmarking services, had commercial value that could be exploited by its competitors. The adjudicator, and later the reviewing court, found that the information at issue was not sufficiently detailed so as to cause commercial harm

if disclosed, and ordered its release. This case, and many others like it, demonstrate that, as with privacy, there can be a delicate balance between the competing interests of protecting confidential commercial information on the one hand and providing an appropriate level of transparency on the other.

It has been argued that some data should be treated as confidential commercial information since their disclosure—alone or in combination with other data—could lead to the reverse engineering of the confidential proprietary decision-making processes or algorithms that were used in the creation of the datasets (see, e.g., Scassa, 2015; *Bertucci v. Royal Bank of Canada*, 2016). Such arguments may create interesting open data challenges for governments. For example, if the release by a city of datasets collected or processed according to proprietary algorithms could lead to the reverse engineering of those algorithms, the third-party company with rights over the algorithms might seek to have that data treated as confidential commercial information that is not subject to release as open data (or under access-to-information legislation). Thus, even where data are not proprietary to a third party, arguments still might be made that their release would lead to the disclosure of other confidential or commercially sensitive information.

2. Crown Copyright and Data Expropriation

Two Canadian court cases suggest that governments may have other tools at their disposal when it comes to balancing the public interest in the disclosure of data as open data with private-sector claims to ownership, confidentiality, or commercial sensitivity, at least in some contexts. In *Geophysical Service Incorporated v. Encana Corporation* (2017), for example, the Alberta Court of Appeal confirmed that a federal regulatory regime which required companies to submit data to a board as part of a licensing process for seismic exploration effectively meant that the federal government acquired the producer's copyright in the data by virtue of the provisions of the *Canada Petroleum Resources Act* (1985) and the related regulatory regimes. Under the legislation, the government was required to keep submitted data confidential for a period of years, after which point they would be published. Notably, the court accepted arguments that companies that used the published data without seeking a licence from Geophysical (the company which supplied it) did not violate that

company's copyright in the data. The court referred to the legislative regime as having a "confiscatory nature" (*Geophysical Incorporated v. Encana Corporation*, 2017, para. 106). The decision confirms that governments, in creating regulatory regimes under statute, have the power to provide for the public interest in access to and use of the data not just by acquiring ownership of them, but by mandating the disclosure of the data, effectively as open data (Scassa, 2021).

In *Keatley Surveying Ltd. v. Teranet* (2019), the issue was whether land surveyors held copyright in their plans of survey, and if so, whether the company operating the provincial land-titles registry system owed them royalties when it charged fees for the reproduction of those plans of survey by users of the digital registry. While confirming that the plans were works protected by copyright law, the Supreme Court of Canada nevertheless ruled that the legislative scheme that governed both the deposit and the publication, through the registry, of the plans of survey resulted in the Government of Ontario acquiring Crown copyright in the documents. As such, it was entitled to control the terms of their dissemination and reproduction. The court ruled that section 12 of the *Copyright Act* (1985) gives the Crown rights in works that are "prepared *or published* by or under the direction or control of Her Majesty or any government department" (emphasis added).

In both *Keatley* and *Geophysical* there were complex legislative regimes in place. This is an important factor in each of these decisions. Although this suggests that a court will require firm legislative grounding before finding copyright to have been usurped by the Crown, the cases demonstrate that government may have latitude to create regimes under which it becomes entitled to publish data that are the subject of proprietary claims by others. This may be a useful tool to manage the public interest in access to some kinds of data.

3. Conclusion

That neither third-party confidential commercial information nor other proprietary third-party data can be released as open data is a limitation that takes on particular importance in the rapidly evolving big data and AI context. Governments may contract with private-sector suppliers for data (e.g., purchasing local cycling data from a fitness-app company; or traffic data from a driving-app company), for data analytics (e.g., software and systems that will analyze input data

and produce analytics for use in predictive policing, or toward public transit arrival times), or for the sensors and systems used to collect and process smart city data. Unless close attention is paid to data ownership or to rights to publish data contracted for by governments, the pool of government data available for open data programs may shrink significantly.

In dealing with confidential commercial data and commercially sensitive data, governments have different options. Some of these are hard-coded in legislation, but it is evident that there are different models to follow. For example, access-to-information laws can create a bright-line rule that excludes confidential commercial information from disclosure, or they can provide latitude for discretionary release where disclosure is in the public interest. Even in the case of bright-line rules, it is possible to have an overarching public-interest override for exceptional cases. In the open data context, it is important to develop guidance and frameworks both for assessing the legitimacy of assertions of confidentiality or commercial sensitivity, but also for guiding discretionary decisions around release.

It is possible to draw lessons from the treatment of personal information in the open data context that can be applied to confidential commercial information or commercially sensitive information. First, governments have the authority to determine that the transparency value or other public interest in some data overrides any confidentiality considerations. There are many examples of where this occurs in the context of personal information. Regardless of its sensitivity, personal information is disclosed by government actors in published court and tribunal decisions, in public registries, in relation to certain activities such as consulting or political donations, and in sunshine laws, to give just a few examples. The Ontario Digital and Data Directive adopts a clear open-by-default approach, while maintaining exceptions for confidential or commercially sensitive data. In dealing with private-sector actors, governments must be more aware of what is being claimed as proprietary or confidential commercial information, and must assess these claims in light of the public interest. They can also implement proactive measures to limit claims that go beyond what is reasonable.

The privacy/open data context also demonstrates how datasets can be subject to anonymization or de-identification techniques in order to permit the release of valuable data without unduly impacting on privacy rights. Similar strategies might be available in the case of

some kinds of confidential or commercially sensitive data. What is required is a process to determine which data within a dataset must remain confidential, and whether the remaining data have value for release as open data.

Recent case law also demonstrates that governments may have other options when it comes to data provided by private-sector actors. Legislative schemes can provide that data submitted to or shared with government are subject to disclosure. Such disclosure may be automatic and routine, or it may be subject to a confidentiality period to provide a kind of limited commercial advantage to offset the cost of creating or compiling the data. The *Keatley* and *Geophysical* cases show that such arrangements are possible, particularly where there is a compelling public interest served by providing access to the data at issue.

References

Borgesius, F. Z., Gray, J., & van Eeechoud, M. (2015). Open data, privacy, and fair information principles: Towards a balancing framework. *Berkeley Technology Law Journal*, *30*(3), 2073–2130.

Government of Canada. (2017). *Open government licence–Canada, 2.0.* Retrieved April 9, 2018, from https://open.canada.ca/en/open-government-licence-canada

Government of Ontario. (2021). *Digital and data directive, 2021*. Retrieved May 21, 2021, from https://www.ontario.ca/page/ontarios-digital-and-data-directive-2021

Green, B., Cunningham, G., Ekblaw, A., Kominers, P., Linzer, A., & Crawford, S. (2017). *Open data privacy*. Berkman Klein Center for Internet & Society Research Publication. Retrieved April 9, 2018, from https://cyber.harvard.edu/publications/2017/02/opendataprivacyplaybook

Hagen, G., Hutchison, C., Lametti, D., Reynolds, G., Scassa, T., & Wilkinson, M. A. (2017). *Canadian intellectual property law: Cases, notes and materials* (2nd ed.). Emond Montgomery Publications.

Information and Privacy Commissioner of Ontario. (2015). *Proactive disclosure of procurement records*. Information and Privacy Commissioner of Ontario. Retrieved April 9, 2018, from https://www.ipc.on.ca/wp-content/uploads/Resources/open-contracting.pdf

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, *29*(4), 258–268.

Johnson, P. A. (2016). Reflecting on the success of open data: How municipal government evaluates their open data programs. *International Journal of E-Planning Research*, *5*(3), 1–12.

Judge, E. F., & Scassa, T. (2010). Intellectual property and the licensing of Canadian government geospatial data: An examination of GeoConnections' recommendations for best practices and template licences. *Canadian Geographer*, *54*(3), 366–374.

Lauriault, T. P., & and Kitchin, R. (2014, April) *A genealogy of data assemblages: tracing the geospatial open access and open data movements in Canada* [Paper presentation]. Data-based Living—Peopling and Placing Big Data Session, Association of American Geographers (AAG) Annual Meeting, Tampa, Fla., United States. Retrieved April 9, 2018, from https://www.slideshare.net/TraceyLauriault/a-genealogy-of-data-assemblages-tracing-the-geospatial-open-access-and-open-data-movements-in-canada

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., & Doshi, E. A. (2013). *Open data: Unlocking innovation and performance with liquid information*. McKinsey Global Institute.

Maynard, C. (2021). *Observations and recommendations from the information commissioner on the government of Canada's review of the access to information regime*. Office of the Information Commissioner of Canada. Retrieved May 21, 2021, from https://www.oic-ci.gc.ca/en/resources/reports-publications/observations-and-recommendations-information-commissioner-review#1

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review*, *57*, 1701.

Open Data Charter. (2015). *Open data charter*. Retrieved April 9, 2018, from https://opendatacharter.net/principles/

Robinson, P. J., & Johnson, P. A. (2016). Civic hackathons: New terrain for citizen-local government interaction? *Urban Planning*, *1*(2), 65–74.

Rocher, L., Hendrickx, J. M., & de Montjoye, Y-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications* 10, 3069. https://doi.org/10.1038/s41467-019-10933-3

Rosenblum, P., & Maples, S. (2009). *Contracts confidential: Ending secret deals in the extractive industries*. Revenue Watch Institute. Retrieved April 9, 2018, from https://www.opensocietyfoundations.org/sites/default/files/contracts_20090915.pdf

Scassa, T. (2010). Geographic information as personal information. *Oxford University Commonwealth Law Journal*, *10*(2), 185–214.

Scassa, T. (2015). Back to the future I: What past privacy findings tell us about the future of big data and privacy [Blog post]. Retrieved April 9, 2018, from http://www.teresascassa.ca/index.php?option=com_k2&view=item&id=180:back-to-the-future-i-what-past-privacy-findings-tell-us-about-the-future-of-big-data-and-privacy&Itemid=80

Scassa, T. (2017). Sharing data in the platform economy: A public interest argument for access to platform data. *University of British Columbia Law Rev*iew, *50*(4), 1017–1071.

Scassa, T. (2018). *Data ownership* (CIGI Paper No. 187). Centre for International Governance Innovation. https://www.cigionline.org/publications/data-ownership.

Scassa, T. (2021). Rights in data, the public interest and international trade law. In I. Bochert & L. A. Winters (Eds.), *Addressing Impediments to Digital Trade*. VoxEU/CEPR. https://voxeu.org/content/addressing-impediments-digital-trade

Scassa, T., & Conroy, A. (2016). Strategies for protecting privacy in open data and proactive disclosure. *Canadian Journal of Law and Technology, 14*, 215–262.

Simperl, E., O'Hara, K., & Gomer, R. (2016). *Analytical report 3: Open data and privacy*. European Data Portal. Retrieved April 9, 2018, from http://www.europeandataportal.eu/

Sweeney, L. (2010). k-Anonymity: A model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems, 10*(5), 557–570.

Treasury Board Secretariat. (2014). *Directive on Open Government.* Retrieved April 9, 2018, from https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=28108

Statutes and Case Law

Access to Information Act, RSC 1985, c A-1.

Agreement on Trade-Related Aspects of Intellectual Property Rights, 15 April 1994, 1869 U.N.T.S. 299. Retrieved April 9, 2018, from https://www.wto.org/english/docs_e/legal_e/27-trips_01_e.htm

Air Atonabee Ltd. v. Canada (Minister of Transport) (1989), 37 Admin. L.R. 245 (F.C.T.D.).

Bertucci v. Royal Bank of Canada (2016), FC 332 (CanLII). Retrieved April 9, 2018, from http://canlii.ca/t/gnzb2

Canada Petroleum Resources Act, RSC 1985, c 36 (2nd Supp.).

Canadian Imperial Bank of Commerce v. Canada (Canadian Human Rights Commission) (2006), FC 443. Retrieved April 9, 2018, from http://canlii.ca/t/1n3p4

Copyright Act, RSC 1985, c C-42.

Freedom of Information and Protection of Privacy Act, RSO 1990, c F.31.

Gartner Inc. v. Ontario (Information and Privacy Commissioner) (2017), ONSC 7181. Retrieved April 9, 2018, from http://canlii.ca/t/hpf5j

Geophysical Service Incorporated v. Encana Corporation (2017), ABCA 125 (CanLII). Retrieved April 9, 2018, from http://canlii.ca/t/h3jnp

Keatley Surveying Ltd. v. Teranet Inc. (2019), SCC 43 (CanLII). Retrieved May 21, 2021, from https://canlii.ca/t/j2kxw

Ottawa Football Club v. Canada (Minister of Fitness and Amateur Sports) (1989), 2 F.C. 480 (T.D.).

R. v. Stewart (1988), 1 SCR 963, 1988 CanLII 86 (SCC). Retrieved April 9, 2018, from http://canlii.ca/t/1ftdt

Simpler, Faster, Better Services Act, 2019, S.O. 2019, c. 7, Sched. 56.

Toronto Real Estate Board v. Commissioner of Competition, 2017 FCA 236 (CanLII), [2018] 3 FCR 563, <https://canlii.ca/t/hp34l>

About the Author

Teresa Scassa is Canada Research Chair in Information Law and Policy at the University of Ottawa, Faculty of Law. She is the author or co-author of several books, including *Digital Commerce in Canada* (LexisNexis, 2020), *Canadian Trademark Law* (LexisNexis, 2015, 2nd ed.), and *Law Beyond Borders* (Irwin Law, 2014); is co-editor of *Artificial Intelligence and the Law in Canada* (LexisNexis, 2021) and *Law and the Sharing Economy* (University of Ottawa Press, 2018); and has written widely in the areas of intellectual-property law, law and technology, and privacy. Scassa is a member of the Canadian Advisory Council on Artificial Intelligence and of the Geothink research partnership. She is also Senior Fellow with the Centre for International Governance Innovation.