Matt Malone

Trade Secrets, Big Data, and the Future of Public Interest Litigation Over Artificial Intelligence in Canada*

Matt Malone**

Abstract

To safeguard big data, many commercial entities deploy legal arguments that these data are trade secrets. Using a Canadian legal backdrop, this article suggests that this argument will hardly be robust, sustainable, or convincing when public interest litigation begins targeting issues of bias and discrimination in artificial intelligence. For now, trade secret litigation is an arena of primarily commercial interests. However, this article suggests that the default view among concerned companies that big data are necessarily trade secrets will be susceptible to attack when public interest litigants turn their sights on matters of bias and discrimination.

Résumé

Pour protéger leurs banques de données, plusieurs entités commerciales présentent des arguments juridiques invoquant que leurs données appartiennent à la catégorie de secrets commerciaux. À l'aide d'un contexte juridique canadien, le présent article suggère que cet argument sera à peine assez robuste, durable ou convainquant lorsqu'un litige d'intérêt public ciblera des questions de partialité et de discrimination dans l'intelligence artificielle. Pour l'instant, le litige en matière de secret commercial demeure un domaine d'intérêts essentiellement commerciaux. Cependant, l'article suggère que la perception par défaut de certaines entreprises concernées à l'effet que leurs données sont nécessairement des secrets commerciaux risque d'être attaquée par les plaideurs d'intérêt public qui tournent leur regard vers les questions de partialité et de discrimination.

Contents

* Submission to the Editor, December 5, 2018.

1 Introduction

For many companies working on artificial intelligence (AI), their greatest asset is data. To justify protecting this asset, these companies often invoke the legal regimes of trade secrets and confidential information. As one experienced intellectual property (IP) lawyer recently commented, "Many companies today count their primary assets and primary worth by data, nothing but data. And the main way that we protect data is with trade secrets."[1] For example, when Tinder was pressed to share data that went into creating each users' so-called desirability ranking, the company stated, "[W]e cannot provide any information that reveals or otherwise compromises all or any part of our proprietary trade secrets or know how."[2] The gist of such arguments is that the data—not the algorithm—are worthy of protection.

In Canada, the legal arguments underpinning the notion that data are susceptible to trade secret protection are largely untested and unstable. Unlike software and algorithms, which manipulate data, and which are clearly works of the mind, big data are not works of the mind *per se*. The question of whether they warrant intellectual property protection is controversial. One view that is increasingly widespread holds that data are a new commodity or resource, tantamount to being the oil of the 21st century, and so a company that collects data is merely exploiting the resource.[3] In Canada, the legal status of such data remains unclear.[4] Of course, there are blurry distinctions—how data are collected determines what data are collected—but the assumption that big data are trade secrets does not rest on clearly established law. At this stage, it is more of an assumption than an argument.

This issue is becoming important as the spotlight turns on tech companies that are facing accusations of promoting political bias[5] and discrimination,[6] inadequately protecting privacy,[7] and fuelling mental health crises,[8] to name to just a few recent controversies. These controversies have been met with calls for greater transparency and openness of data—clashing with companies' desires to hoard data in their exclusive possession. A warning sign of these future skirmishes is the recent dispute between the City of Seattle and a consortium of Uber and Lyft, in a case filed in the Supreme Court of Washington, which raised a significant allegation of racial discrimination related to the manipulation of data.[9] The drama revolved, in part, around a demand to make these data available to public authorities for oversight. Unsurprisingly, Uber and Lyft both argued that the data were trade secrets. As the case shows, public access to data will increasingly become a justiciable question, and the issue of whether they are trade secrets will become a central issue. It will be at the heart of many future disputes over data transparency, sharing, and openness.

2 The Law of Trade Secrets in Canada

Trade secret law protects sensitive business information that acquires value from not being known to the public. The classic test for the determination of trade secrets is inexorably bound up with the duty of confidence, from which it remains, for practical purposes, indistinguishable. The most binding formulation of this duty appeared in *Lac Minerals Ltd v International Corona Resources Ltd*,[10] where the Supreme Court of Canada examined the nature of duties owed by executives of a large mining company to a development-stage junior competitor whose lucrative secret business information the larger company misappropriated. The court affirmed the following test for the breach of confidence: (1) the existence of confidential information, (2) its communication in confidence, and (3) its misuse by the party to whom it was communicated. Undoubtedly, the hardest component of this test for a plaintiff to satisfy is the third prong.

Historically, the classic formulation of a trade secret case involved theft by a departing employee of sensitive information such as a recipe or a customer list created by the ex-employer. The general principle was that the law would not countenance giving the ex-employee unfair advantage for stealing something that the ex-employer had created and sought to keep secret. Part of a broader commercial morality, these principles over time proved adaptable in their extension to new fields of technology—including Listerine,[11] stock-trading platforms, and even light detection and ranging (LIDAR) sensor technology used by autonomous driving vehicles,[12] to name just a few prominent trade secret cases.

Defining big data as trade secrets, though, is somewhat trickier. Although the definition given above shows that subject matter is usually a trade secret when the party claiming protection has "used his brain," as one English judge put it more than 70 years ago,[13] this argument is less persuasive when applied to data. This is because, in many cases, algorithms or software collect or

1 J Pooley, 2018, <www.pooley.com>.
2 C Ip, "Who Controls Your Data?" (2018), *Engadget* (blog); A Carr, "I Found Out My Secret Internal Tinder Rating and Now I Wish I Hadn't" (2016), *Fast Company* (blog).
3 Centre for International Governance Innovation (CIGI), *Data Governance in the Digital Age*, Special Report (Waterloo, Ont: CIGI, 2018) [CIGI].
4 T Scassa, "Data Ownership" (4 September 2018) CIGI Paper No 187; University of Ottawa Faculty of Law Working Paper No 2018-26.
5 R Waters, "Google Defends Search Algorithms Against Bias Claims", *Financial Times* (24 September 2018).
6 E Siegel, "When Machines and Data Promote Blatant Discrimination", *San Francisco Chronicle* (21 September 2018).
7 T Spangler, "Facebook Under Fire: How Privacy Crisis Could Change Big Data Forever", *Variety* (3 April 2018).
8 W Ghonim & J Rashbass, "It's Time to End the Secrecy and Opacity of Social Media", *The Washington Post* (31 October 2017) [Ghonim & Rashbass].
9 D Gutman, "Uber and Lyft May Have to Disclose Seattle Data They Claim Secret, Supreme Court Rules", *The Seattle Times* (31 May 2018); *Lyft Inc v City of Seattle*, file no 94026-6 (Wash Sup Ct 2018).
10 *Lac Minerals Ltd v International Corona Resources Ltd*, [1989] 2 SCR 574.
11 *Warner-Lambert Pharm Co v John J Reynolds Inc*, 178 F Supp 655 (SDNY 1959).
12 A Marshall, "Uber and Waymo Abruptly Settle for $245 Million", *Wired* (9 February 2018).
13 *Saltman Engineering Co v Campbell Engineering Co Ltd*, [1948] 65 RPC 203.

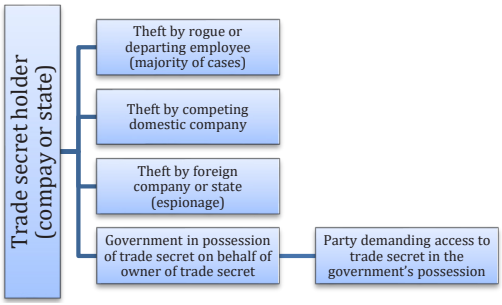
"scrape" data and, in other cases, individuals self-populate data when they interact with social media or other Internet of Things (IoT) applications. These interactions yield data collections, but the question of who created the data is epistemological: Were the data created by the software designers who conceived of technologies that captured the data? Or do legal rights inhere in the data of individuals who input the data into those technologies?

Of course, it may be argued that it is impossible to separate the method of collecting data from its content, and some academics argue that big data would fall into the category of trade secrets for this reason.[14] But Canadian judges in a few cases have split hairs on this question and come to opposite conclusions. For example, in a recent Alberta decision handed down by the Office of the Information and Privacy Commissioner, a party sought access from a provincial agency to a third party's data on gas and oil drilling costs in particular areas of the province.[15] The judge held that the data, separated from the techniques used to collect them, were a "compilation and/or product" that were therefore trade secrets. Conversely, in an Ontario decision issued by the Information and Privacy Commissioner, the Limestone District School Board possessed survey data on school improvement that a party sought to access under the *Municipal Freedom of Information and Protection of Privacy Act*.[16] The judge held that while the questionnaire that was used to collect the survey data "might reveal a unique method, formula, pattern or compilation of information" and hence a trade secret, the survey data themselves were not trade secrets.

The commercial advantages of seeking trade secret protection for data over other forms of intellectual property are many. Unlike a patent, there is minimal cost involved in claiming trade secret protection. Patents also require time investment and burdensome renewal procedures, whereas trade secret protection starts immediately and requires minimum safeguards. And unlike the protection offered by a patent, which is time-capped, trade secret protection can last forever. It can also cover a range of subject matter that is not patentable at all. Thus, at its core, unlike areas of "hard" IP that require formal registration, the "soft" area of trade secret protection is triggered by conduct: if a party acts like something is a trade secret, and protects it as such, then it is arguably a trade secret.

The operative word here is "arguably." Trade secret law is tainted with imprecision, because the threshold question in every trade secret case is essentially: "What is a trade secret?" Textbook definitions and previous judgments can only guide such a determination. The lack of precision in the legal definition of a trade secret, however, gets at an essential element of trade secret law—once the item is no longer secret, its protection vanishes. Trade secret law, then, usually only goes to the courts when there is a dispute over the misappropriation of the secret. Litigation arises as part of a commercial strategy—defensive or offensive—by one actor to prevent another actor from gaining competitive or economic advantage. Accordingly, most trade secret cases follow one of the scenarios illustrated in figure 1.

Figure 1



As shown in figure 1, most disputes involve an (ex-)employer and (ex-)employee.[17] Some practitioners argue that some 90 percent of trade secret cases involve these scenarios.[18] As for the other types of disputes, Canada has seen disputes between commercial actors, such as the high-profile Air Canada–WestJet spying lawsuit and instances of espionage by foreign states and companies—though it lacks sufficient espionage provisions to crack down on them, and has sought to resolve them through diplomatic efforts.[19] Finally, Canada has also often witnessed cases involving the government's disposition of a third party's trade secrets (in the mold of the previously cited Alberta and Ontario cases).

The last category is the most relevant one to litigation concerning big data as trade secrets in Canada. These disputes involve a government entity possessing trade secrets that belong to a third party, which an outside actor seeks to obtain through an access to information request. The highest-stakes versions of these disputes emerge when the subject matter in question is pharmaceutical companies' regulatory and clinical data in the possession of Health Canada.

A recent example of this type of dispute is the Federal Court case *Doshi v Canada*,[20] which revolved around the confidentiality of datasets in clinical trials. The case concerned a researcher at the

14 M Mattioli, "Disclosing Big Data" (2014) 99:2 Minn L Rev 535.
15 Alberta Energy, Alberta Office of the Information and Privacy Commissioner, Order F2015-15, case file no F6260.
16 *Limestone District School Board (Re)*, 2013 CanLII 77839 (ON IPC).
17 DS Levine & CB Seaman, "The DTSA at One: An Empirical Study of the First Year of Litigation Under the Defend Trade Secrets Act" (2018) 53:1 Wake Forest L Rev 105.
18 M Klapow, J Milewski & W Pellett, "How Courts Approach Trade Secret Identification" (2018) *Law 360* (blog).
19 I Austen, "WestJet Settles Spy Case with Rival Air Canada—Business—International Herald Tribune", *The New York Times* (30 May 2006); N Fraser, "Canadian Law on Industrial Espionage Needed", *Law Times* (6 March 2006); R Fife & S Chase, "Canada and China Strike Corporate Hacking Deal", *The Globe and Mail* (25 June 2017).
20 *Doshi v Canada (Attorney General)*, 2018 FC 710 [*Doshi*].

University of Maryland who was seeking access to clinical data related to HPV vaccines and neuraminidase inhibitors under a legal exception for researchers gaining access to data for "the protection or promotion of human health or the safety of the public."[21] Health Canada, which under the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) and the North American Free Trade Agreement (NAFTA) was obliged to protect trade secrets, would provide the information only if the researcher agreed to sign a confidentiality agreement. After the researcher did not sign the agreement, Health Canada refused to share the information. The researcher submitted a request for judicial review. Drawing a distinction between language in TRIPS and NAFTA that differentiated between "undisclosed information," "trade secrets," and "data," Justice Grammond found that the data should be shared in light of the public interest exception. The dispute presented an interesting example of how the defence of trade secrets does not hold up against the needs and demands of the public interest.

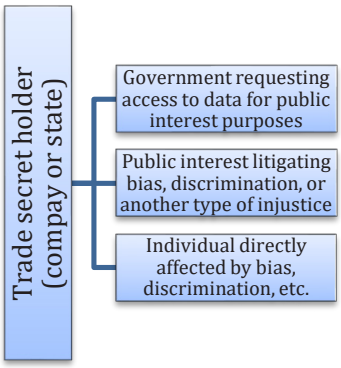
Although Justice Grammond rooted his analysis in a subtle distinction between the statutory definitions of "trade secrets" and "data" in TRIPS and NAFTA, the rejection of a trade secret classification for the data was undoubtedly motivated by a public interest argument. The notion of "public interest" was directly invoked only once in the case—in the final paragraph of the decision, where the parties agreed not to seek costs—but it courses throughout the decision. Justice Grammond characterized the competing interests of drug-related legislation as "protecting the health and safety of the public and promoting the economic interests of pharmaceutical companies." As he noted, public disclosure of information otherwise susceptible to trade secret protection may be "beneficial" where "[t]here are concerns that the conduct of those tests may be *biased*."[22]

As his ultimate determination in *Doshi* revealed, the inherent difficulties associated with defining business secrets mean that deploying such an argument to shield disclosure does not hold up against the needs and demands of the public. Secrecy itself is a term inflected by cultural and societal values, which hinge on public interest. Therefore, the value of trade secrets to the public will always exert influence on that determination. But while trade secret disputes between commercial actors usually involve a relevant sum, when the public interest is at stake, the value of the trade secret becomes less persuasive.

Admittedly, disputes like the *Doshi* case are still rare. Most trade secret disputes in Canada do not involve a confrontation between public interest calls to make data more transparent and companies seeking to hoard data. Calls for open data and data transparency, while popular, are rarely litigated and do not appear to fit easily into the modalities of trade secret litigation presented above. Not faced with any imminent litigation or threat thereof, companies have not had to worry about handing over their data, let alone fine-tuning legal arguments that prevent them from being required to do so. What this shows is that, despite the uproar about the deleterious side effects of the current unilateral data-hoarding approach, the paradigm of data hoarding has so far gone mostly unchallenged in Canada.

When public interest actors begin to target these companies over such practices, however, trade secret litigation will start to look a lot different. Looking beyond the old types of trade secret disputes, future disputes over data will arise regarding public interest issues concerning bias, discrimination, mental health, and so on, and involve areas of law like the *Charter of Rights and Freedoms* and privacy law. When such issues are litigated, the defensive argument that data are trade secrets will confront public interest head on. As opposed to the scenarios shown in figure 1, disputes over explosive issues like racially or gender-motivated bias and discrimination in AI are more likely to be structured as shown in figure 2.

Figure 2



If the outcome of *Doshi* is a harbinger of future results, the public interest will prove very effective at destroying the argument that data are trade secrets. That should be good news for public interest litigants in Canada who want to take on companies trying to safeguard their data. Nonetheless, there remain significant hurdles in organizing such litigation. Until now, calls for greater transparency have lacked bite because there has been an inordinate focus on the benefits that society can reap from rendering data more available—rather than on how greater transparency may actually be achieved. Suggestions such as developing a "standardized public interest API that provides a detailed overview of the information" get at what an ideal open data sphere may look like, but they do not show us how to get there without applying pressure on commercial entities.[23] Even within the fascinating legal work on data governance that is being done, there is little discussion about the strategies that would actually compel companies to share their data.[24] For now, the siren is sounding. But as the history of trade secret litigation involving data in Canada shows, there is reason to believe that many assumptions about the law affording security against disclosure are simply wrong.

21 *Food and Drugs Act*, RSC 1985, c F-27, s 21.1(3).
22 *Doshi*, *supra* note 20 at para 11 (emphasis added).
23 Ghonim & Rashbass, *supra* note 8.
24 CIGI, *supra* note 3.