Comments on the OPC Consultation on Artificial Intelligence

February 19, 2020

Introduction

On January 28, 2020, the Office of the Privacy Commissioner of Canada (OPC) published its Consultation on the OPC's Proposals for ensuring appropriate regulation of *artificial intelligence* (Consultation Paper). The Consultation Paper sets out several proposals for how the federal *Personal Information Protection and Electronic Documents Act* (PIPEDA) could be reformed, in the words of the OPC, "in order to bolster privacy protection and achieve responsible innovation in a digital era that involves artificial intelligence (AI) systems." The document also invites privacy experts to validate the OPC's understanding of how privacy principles should apply and whether its proposals would be consistent with the responsible development and deployment of artificial intelligence, calling for responses to be submitted by March 13, 2020.

The Consultation Paper considers the perspectives of other bodies that have treated the issues raised by the various proposals at length, including the OECD, the IEEE and the UK Information Commissioner's Office among others. This makes the document substantial, and commenting on the Consultation Paper in its entirety is not feasible in a short post. In consequence, this bulletin will provide critical commentary on a few of the more salient points of interest.

There is no question that the recent convergence of cheap, on-demand computing resources, very large data collections and the development of machine learning platforms makes it timely to consider legal reforms that better address the promise and the risks surrounding the latest wave of developments in AI.

That said, any consideration of this topic must begin with a caveat: the term "artificial intelligence" has a long history in computer science, cognitive science and philosophy, and its meaning has become rather elastic.[1] This can be useful for marketing but hinders legal analysis.

Defining AI

The first proposal provides a case in point. It considers whether reforms to PIPEDA should incorporate a definition of AI within the law that would serve to clarify which legal rules would apply only to it, and proposes the definition from the 2019 OECD *Principles on Artificial Intelligence* to which Canada is signatory as a possible contender:

> "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy".

The proposed definition, however well intentioned, is so broad as to be of little value for the intended purpose. Spam detection software falls squarely within it, yet that is clearly not what the OPC wishes to address with the proposed reforms. Decision-making software that takes its cues from a random noise generator could also fit within this scheme.[2]

If your reaction to the latter example is that random value selection is not a decision-making process, proceed with caution: a brief glance into scholarly literature about what constitutes a decision reveals discussion that typically imports notions of *reasoning*, *considering*, *thinking* and *choosing*,[3] all of which are questionable concepts to deploy in the context of discrete state machines.

There is another problem to consider: there are many, many definitions of AI, each having their own merits and their own issues.[4] Reading several of them in quick succession can lead to confusion as to whether the goal of such definitions is to describe systems that satisfy certain criteria in relation to cognitive processes, or just externally observable behaviour. Moreover, some definitions introduce terms that lead to thornier issues similar to those alluded to above, such as *perceiving*,[5] *learning*,[6] or *agency*.[7]

It is difficult to avoid the conclusion that AI is an aspirational term that means fundamentally different things to different people. That is not a good foundation upon which to build a definition that will be enshrined in law. To the extent that legislators are only interested in drafting a definition that addresses the latest wave of technological developments that permit more sophisticated predictions and decision-making capacities than earlier iterations of computing machinery were capable of the slogan "it's only AI until you understand it; then it's just software" comes to mind. Although somewhat glib and condescending, it also contains a kernel of truth.

All AI currently in play is "narrow" AI,[8] much of it is also "weak" AI,[9] and frankly none of it is intelligent (said while acknowledging that human intelligence is not necessarily the only kind of system we could recognize as intelligent).[10] The OPC appears to endorse, through quotation, the view that many of the machines we will interact with in the future will be "more akin to agents than mere devices",[11] but note the use of the word 'akin'. If we ever manage to create machines that we genuinely regard as having agency and not just the appearance of agency, the law will require significant reform, and not just in the domain of privacy and data protection. We are not there yet. As such, for the time being, AI systems could be "governed by the same rules as other forms of processing", as the OPC puts it.

A rights-based approach

The Consultation Paper also proposes the introduction of a rights-based approach, noting that it would be "consistent with many international instruments, including the GDPR, which has incorporated a human rights-based approach to privacy within the EU's data protection legislation".

Adoption of this proposal would likely allay some of the concerns that have arisen around the widespread deployment of AI systems in circumstances where those systems "make decisions for and about us with little to no human involvement".

In considering this proposal, the most important question to ask is a very general policy question: how may we best arrange our institutions that govern privacy and data protection in a way that protects individual privacy while allowing us to reap the benefits that this technology offers?

The OPC proposal, to reimagine PIPEDA as a fundamentally rights-based legislative instrument, could be seen as a departure from the current legal framework that seeks balance by recognizing both individual privacy rights and the needs of organizations.[12] The OPC has mentioned that balance on numerous occasions, most recently in its 2019 annual report to Parliament.[13] In that annual report, however, the OPC rejects what it sees as an implication arising from this discourse to the effect that privacy rights and economic interests are engaged in a zero-sum game, noting, "a rights-based statute would serve to support responsible innovation by promoting trust in government and commercial activities."[14]

The OPC may be correct. No matter what approach or framework is settled upon, the question is whether it will protect individual rights while still supporting explorations of this technology that can lead to public benefit, economic or otherwise. There is no reason that the OPC's proposal would fail in principle, but it may be challenging to adopt a rights-based framework in a way that will provide sufficient latitude in that regard. The challenge could be met, in part, by providing alternative legal grounds for processing in certain circumstances, as suggested by one of the Paper's later proposals, which states that alternative grounds "should be available in instances where obtaining meaningful consent is not possible." While that sounds like a practical solution, to the extent that the OPC wishes to put robust limits around the invocation of alternative legal grounds, it puts a great deal of pressure on the concept of meaningful consent. The next section considers whether that notion can take the strain.

Transparency, explainability, and interpretability

Supporting the OPC's default position that organizations using AI should still obtain meaningful consent where possible, the Consultation Paper includes a proposal to "[p]rovide individuals with a right to explanation and increased transparency when they interact with, or are subject to, automated processing." To the extent that anxiety over the use of AI systems for automated decision-making arises in part because (for most members of the public) they are mysterious black boxes, it is worth making this a focus of attention.

The OPC notes, as presently framed, the transparency principle lacks specificity. Better articulation of the transparency principle could greatly assist both individuals and organizations, and an explainability component could further assist in that regard, but

only if the new law provides robust guidance on how transparency and explainability should play out in practice.

There is a good deal of uncertainty on the part of organizations as to how much explanation is appropriate and/or necessary when dealing with highly sophisticated processing systems, of which AI is just one example, particularly where such disclosures might reveal trade secrets. Having more explicit direction in the law could assist organizations in understanding their obligations and how those interact with their rights to maintain confidentiality around certain aspects of their systems, and if the new provisions are carefully fashioned, the outcome could be better individual understanding of how these systems "work" in ways that are pertinent to providing meaningful consent.

The challenge here should not be underestimated, however, given that the most prominent target for the AI reforms are the most sophisticated of these systems, the "deep learning" neural network architectures. The internal workings of these AI systems are notoriously opaque, even to their designers, and may not be explainable in the sense desired by the OPC.

Which leads us to a useful distinction that is sometimes made between explainability and interpretability.[15] Interpretability is concerned with comprehending what the target of interpretation has done with a given input, or might have done with a different input. Explainability is concerned with providing the reasons for a given output.

Many systems that we interact with every day are interpretable to us, but not necessarily explainable: a mobile phone, a car, an elevator. For most people, such systems are black boxes. Through experience, individuals will come to associate a variety of inputs with outputs, reactions or responses, and can also make successful predictions about how one of these systems would react to a certain input (even if that particular input had never been made by that individual). For such individuals, such systems are interpretable.

Yet, faced with an unexpected output, individuals who have learned only how to interpret a system may be unable to explain the result because they do not truly understand the system. No behaviour of a system that is fully explainable will be unexpected, apart from malfunctions. Even if a system is explainable in that sense to an expert, however, it may not be explainable to the average individual. That is why we typically rely on interpretability: we skip the (many) details that we just would not understand anyway.

Does the OPC seek interpretability or explainability? The Consultation Paper does not invoke this distinction. Some of the OPC's comments suggest that it is trying to come to grips with some aspects of it, but those remarks also suggest that the office does not entirely understand the nature of the beast that it is trying to wrangle.

The OPC states that individuals should be provided with "the reasoning underlying any automated processing of their data, and the consequences of such reasoning for their rights and interests". This suggests that the OPC is interested in requiring explainability. The OPC also states that it might support the notion of public filings for algorithms, and under another proposal, the OPC also seeks a requirement for algorithmic traceability. This suggests that the OPC imagines that the mechanics of all AI systems are amenable

to an algorithmic reduction that makes them explainable, that the “reasoning” of these systems can be communicated in a meaningful way.

A true algorithmic trace of a deep learning system, moving stepwise from input to output, may be recoverable from examination of the weighted nodes and their interconnections; but the “reasoning” behind each step, and the sheer number of steps, would yield an algorithm that is no more comprehensible to regulators and individuals than it is to the system’s designers. The patterns of interactions created by those nodes and interconnections are abstracts, complex and use clusters of factors that make "no intuitive or theoretical sense."[16] Providing this information to individuals will not create the conditions for meaningful consent.

In fact, the question as to whether to provide full explanations or just enough information to make a system interpretable for the average individual predates the existence of automated decision-making systems. With the advent of deep learning AI, however, the problem is thrown into sharp relief.

As such, while it is laudable for the OPC to be devoting attention to matters of transparency and explainability, in order to provide a practical legal framework it will need to give far more attention to this problem than it may have anticipated.

The right to object

The Consultation Paper also considers a proposal to provide a right to object to automated decision-making and not to be subject to decisions based solely on automated processing, subject to certain exceptions. Such a right is worth considering provided the exceptions are broadly drafted. The GDPR, as the Consultation Paper notes, provides exceptions when an automated decision is necessary for a contract; authorized by law; or where explicit consent is obtained.

This is a reasonable approach. Although at present we may be skeptical as to the quality of decisions provided by these systems, we may eventually come to place more trust in the decisions they deliver than those of humans, in certain circumstances. The discourse in autonomous vehicles provides an interesting example: the technology has shown enough promise that regulators, municipalities, and insurers are considering a future in which there could be fewer accidents and more efficient traffic flows where in-vehicle automated systems make the decisions. That might ultimately lead to a future in which we would want to curtail the right of individuals to intervene in the driving process, and we may even come to expect that reasonable people would not try to drive manually. Any reforms in PIPEDA that import a right to object to automated decision-making should be drafted to accommodate shifts in reasonable expectations and public policy.

Conclusion

Reform of Canada’s privacy laws is needed, and some of that reform should be crafted with AI in mind. Based on what the Consultation Paper discloses, however, it is not feasible to validate completely those of the OPC’s proposals that were discussed in this bulletin. While there is merit in those proposals, attempting to create a special regime to address AI directly (however defined) at this stage of its development would be

premature; we have only inklings of how the latest wave of developments will ultimately play out. In the face of such uncertainty, we should maintain the flexibility that a law of general application can provide.

[1] Coined by the computer scientist John McCarthy in 1955 in his proposal for the first conference organized to discuss the subject, the so-called “Dartmouth Conference” of 1956, the aim of which McCarthy framed as follows in the 1955 proposal circulated to interested scientists: “An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.” See John McCarthy, Marvin Minsky, Nathan Rochester, and Claude Shannon, “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence” (1955).

[2] One can imagine software that provides a virtual roulette wheel, the human-defined objective of which is to ensure that the long-term average outcome of spins of the wheel approaches a certain ratio of winners to losers. The software makes decisions as to who is a winner or a loser for a given spin by consulting the random noise generator to determine spin outcome. Its decision-making procedure is even fairly autonomous: each individual user provides input in the form of selecting a number (or color, parity, etc.), but the system itself operates and provides output without further human intervention.

[3] See *e.g.* Samuel Eilon, “What is a Decision?,” *Management Science*, Vol. 16, No. 4, Application Series (Dec., 1969), pp. B172-B189.

[4] See *e.g.* Russell, Stuart J.; Norvig, Peter (2009). Artificial Intelligence: A Modern Approach (3rd ed.). Upper Saddle River, New Jersey: Prentice Hall. ISBN 978-0-13-604259-4, pp. 1-2.

[5] Poole, David; Mackworth, Alan; Goebel, Randy (1998). Computational Intelligence: A Logical Approach. New York: Oxford University Press. ISBN 978-0-19-510270-3, p.1.

[6] Bellman, R. E. (1978). "An Introduction to Artificial Intelligence: Can Computers Think? Boyd & Fraser Publishing Company".

[7] Poole, D., Mackworth, A. K., and Goebel, R. (1998). "Computational intelligence: A logical approach". Oxford University Press.

[8] That is, designed to perform a narrow set of tasks.

[9] See John Searle, "Minds, Brains and Programs", Behavioral and Brain Sciences, 3 (3)1980: 417–457, doi:10.1017/S0140525X00005756 for the original formulations of “strong” and “weak” AI.

[10] See *e.g.* James Vincent, “This is when AI’s top researchers think artificial general intelligence will be achieved”, *The Verge*, 27 November 2018; Andrey Kurenkov, “AlphaGo Zero Is Not A Sign of Imminent Human-Level AI”, Skynet Today, 30 March 2018; Martin Ford, *Architects of Intelligence* (Birmingham: Packt Publishing Ltd.), 2018.

[11] Ian Kerr, “Robots and Artificial Intelligence in Health Care,” Canadian Health Law and Policy, 5th edition, 2017, p.279.

[12] *Personal Information Protection and Electronic Documents Act*, SC 2000, c 5, s. 3.

[13] "*2018-2019 Annual Report to Parliament on the Privacy Act and the Personal Information Protection and Electronic Documents Act*".

[14] *Ibid*.

[15] Leilani H. Gllpin et al., “Explaining Explanations: An Overview of Interpretability of Machine Learning,” February 3, 2019.

[16] Mark, MacCarthy, "How to address new privacy issues raised by artificial intelligence and machine learning," *Brookings*, April 1, 2019.

By

Max Jarvie

Expertise

Cybersecurity, Privacy & Data Protection, Technology, Artificial Intelligence (AI)

BLG | Canada’s Law Firm

As the largest, truly full-service Canadian law firm, Borden Ladner Gervais LLP (BLG) delivers practical legal advice for domestic and international clients across more practices and industries than any Canadian firm. With over 725 lawyers, intellectual property agents and other professionals, BLG serves the legal needs of businesses and institutions across Canada and beyond – from M&A and capital markets, to disputes, financing, and trademark & patent registration.

blg.com

BLG Offices

Calgary

Centennial Place, East Tower
520 3rd Avenue S.W.
Calgary, AB, Canada
T2P 0R3

T 403.232.9500
F 403.266.1395

Ottawa

World Exchange Plaza
100 Queen Street
Ottawa, ON, Canada
K1P 1J9

T 613.237.5160
F 613.230.8842

Vancouver

1200 Waterfront Centre
200 Burrard Street
Vancouver, BC, Canada
V7X 1T2

T 604.687.5744
F 604.687.1415

Montréal

1000 De La Gauchetière Street West
Suite 900
Montréal, QC, Canada
H3B 5H4

T 514.954.2555
F 514.879.9015

Toronto

Bay Adelaide Centre, East Tower
22 Adelaide Street West
Toronto, ON, Canada
M5H 4E3

T 416.367.6000
F 416.367.6749