

## Towards a Better Public Sector AI Strategy

1500 words

Will Alexander - [walex025@uottawa.ca](mailto:walex025@uottawa.ca)

University of Ottawa - Master of Arts Public Administration

Supervisor: Dr. David Brown - [DavidCG.Brown@uottawa.ca](mailto:DavidCG.Brown@uottawa.ca)

In the midst of the government of Canada's evolving digital environment, Artificial Intelligence (AI) has emerged as a new potential tool for use in the public sector. AI systems, broadly defined, are machines capable of carrying out a particular task that previously could only be done by human beings — learning and reasoning, among others. Interest is growing rapidly in leveraging their potential, but the Canadian government lacks a comprehensive and adequately technical plan to deal with the upturn of emerging AI applications. This essay will argue that, despite increasing interest in “high tech solutions,” the government's current strategy to manage AI-public sector integration is insufficient in its capacity to adequately address AI implementation; it will then explore challenges unique to AI; and finally propose a strategic path forward for the Canadian public service.

### **AI and The Canadian Public Sector**

The Canadian government's AI strategy is well-intentioned, but is limited by a deficit in technical expertise. Post-New Public Management governance has been broadly characterized by evolving IT systems (Dunleavy et al. 2006). The most recent Clerk's Report stressed the importance of “using technology to do things better” (Charette 2022), and in the ‘Beyond 2020’ plan designed to revitalize the public service, one of the three primary areas of focus specifies that the government will look to “explore technology and tools” to increase public service effectiveness (PCO 2020). The Government of Canada had laid the groundwork for future work in this sector. In 2021 the government invested \$208 million in research over ten years through several institutes collectively known as the Canadian Institute for Advanced Research (CIFAR) (ISED 2022).

The Canadian Treasury Board Secretariat (TBS) has put forth guidelines for designing public sector AI systems (TBS 2021); however, their methods have limited understanding of the technology that they plan to regulate. Particularly telling of the Canadian federal government's discomfort and unfamiliarity with this area of technology is the TBS's Algorithmic Impact Assessment (AIA) Tool which is composed of superficial yes/no questions, and which only four departments or agencies had filled out as of 2022 (Longo 2022, p.387).

The government's current AI strategy is crystallized in the "Directive on Automated Decision Making" — a subsection of TBS's "Policy on Service and Digital" designed to regulate public sector AI (TBS 2019). In CIFAR's most recent strategy report, the Advisory Council on AI is credited as the public body responsible for informing Canada's AI strategy (Chowdhury et al. 2020).

The Advisory Council has repeatedly proposed solutions to government, but their ideas have manifested themselves in highly reductive and ambiguous forms in policy. The Council's 2021 Annual Report made frequent reference to "advising leadership on risks and challenges" (ISED 2021). The Advisory Council met sporadically in 2022 — in February some members "suggested that the Council should explore regulatory models for AI." The meeting notes from May cite the need for "an agile regulatory infrastructure that can be responsive to complex sets of challenges." Finally, in September of the same year, a meeting was called to discuss an impending review of the Directive on Automated Decision Making (ISED 2022). In this review, twelve amendments were proposed. Many were procedural, but a few appeared to herald an improvement in the comprehensiveness of Canada's AI strategy. Of note: Expanded AIA questions establishing explanation criteria, expanded bias testing, and tracing measures for data produced by AI systems ("Third Review..." 2022).

Despite this work, the Directive itself has still not evolved meaningfully since it first took effect in 2019. Appendix C of the Directive contains impact assessment indicators. The importance of a “human-in-the-loop,” one of the most common mitigation measures suggested by experts (Wirtz 2020), has been added but the Directive’s detail is limited to the stipulation that “final decisions must be made by a human.” Likewise, “explanation requirement” is present as a criterion, but only for instances where services are denied and without mention of “explainable AI,” a critical ethical consideration highlighted primarily by computer scientists (Newell 2015; Yampolskiy 2020).

Finally, to achieve the highest standard available to public sector AI the Directive requires that systems must receive Treasury Board approval, but no information is provided on how TB will conduct evaluations or what expertise they will rely on beyond an ambiguous list of experts up for “Peer Review” (TBS 2019). After five years of development and three reviews there is still no sign of an “agile regulatory infrastructure” for managing public sector AI, and no evidence that expert suggestions have amounted to anything more than an elaborate paper exercise.

## **AI Implementation**

TBS is dealing with the early stages of a technology implementation problem. It is no secret that the Canadian Government has struggled with modernizing its IT systems in recent years. The NPM era reorganized government into “discrete corporate hierarchies,” siloing departments, and increasing reliance on private sector contracts (Dunleavy et al. 2006, p. 227). Many public sector academics remark that the modern public service is ill-equipped to take on the process of thoroughly evaluating its technological backbone (Clarke et al. 2017; Valle-Cruz

2020; Wirtz 2020). The private sector has not stepped up to lead on matters of ethics and risk mitigation, leaving the public sector at an early disadvantage to take on the challenges of AI (Fatima et al. 2020). There is also a technical knowledge gap in the public service — and more concerningly, there is evidence of a disconnect between AI researchers at CIFAR and the Advisory Council and their counterparts in the general public service.

This disconnect is particularly pronounced in the case of AI for several reasons. Generally speaking, there is an inverse relationship between an AI system’s explainability and its raw power. The most effective methods of machine learning — neural networks, classification trees, and vector machines — can be indecipherable even to professional computer scientists (let alone policy experts) (Yampolskiy 2020). Likewise, for a successful human-in-the-loop model, the eponymous human must be capable of understanding stakeholder concerns in addition to the technical parameters of their system (Wirtz 2020). Expertise in both machine learning and public policy is required.

Finally, risk is often invisible when looking at an algorithm’s architecture — training data and poor parameterization alone can automate systemic discrimination, as was the case with a notorious recidivism-assessing algorithm that inherited discriminatory inclinations from its training data (Angwin et al. 2016). Bias alone is not a viable indicator to tell if an algorithm will be problematic; Hildebrandt remarks that bias is “a generative and inevitable precondition” for AI (Hildebrandt 2021), meaning that AI favours certain patterns of data over others by design, and that an ill-defined review process is not enough to eliminate risk. The “bias testing,” “data tracing,” and ‘explainability’ amendments accepted by TBS were extremely ambitious undertakings that have yet to be meaningfully translated to the Directive.

These three factors combined create a pernicious environment for the Canadian Government and public. If the TB is given ultimate authority for greenlighting algorithms, there is a risk that government falls victim to a phenomenon that Haque describes as “faith in technological rationality” in which, lacking the technical understanding to critically evaluate or implement, decision-makers in the public sector follow the advice of technical experts blindly (Haque 2015).

### **A Way Forward**

The impending challenge of public sector AI is the upshot of complex subject matter and a public-sector technical deficit. AI is already involved in increasingly sophisticated public sector functions, and has begun to embed itself deeply into government operations worldwide (Engstrom and Ho 2020). However, the government of Canada is not yet on a path where adoption is a safe or realistic possibility. Below are two recommendations to course-correct the federal public sector.

- 1. Develop a detailed and comprehensive AI implementation and regulation strategy.**

In a 2022 report from CIFAR, Raji laments the ‘over the wall’ paradigm of IT design, where tools are developed by engineers and tossed ‘over a wall’ to users without consideration (Raji 2022). This approach limits the latitude that client organizations can take with their technology and precipitates problems — particularly evident in strategies where the government has opted for off-the-shelf products (as was the case with Phoenix) or decided to outsource everything (as was the case with ArriveCan) (OAG 2018; Leblanc 2023). To avoid these woes with AI, the government should produce a comprehensive strategy to integrate AI into the public sector. The current Directive

guidelines fall short as mitigation measures, but there is ample good advice circulating among organizations that the government has already funded. Channeling this knowledge into a central pillar of codified practice would be a vast improvement from current measures.

2. **Reestablish in-house technical expertise.** While NPM did away with much of the public sector's reflexive IT capacity, public sector IT needs have only grown. Private sector contracts are expensive and constricting, and private sector experts often fail to fully address the needs of the government. By making public servants ambassadors between the understandable limits of AI and the judgement and direction of officials, it may eventually be possible to reliably translate public sector values into AI system design.

## Bibliography

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine bias: There's software used across the country to predict future criminals. and it's biased against blacks. ProPublica, May 23. Retrieved October 30th, 2022.

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

Charette, Janice. *29th Annual Report to the Prime Minister on the Public Service of Canada*. 2022. <https://www.canada.ca/en/privy-council/corporate/transparency/annual-report-prime-minister-public-service/29th.html>

Chowdhury, N., Hakim, Z., Kim, T., Taylor, N. A., Remennik, T., Rogers, S., Strome, E., & Wallis, J. (2020, October). Pan-Canadian AI strategy. CIFAR. Retrieved January 29, 2023, from <https://cifar.ca/ai/>

Clarke, Amanda, et al. "Understanding Governance in the Digital Era: An Agenda for Public Administration Research in Canada." *Canadian Public Administration*, vol. 60, no. 4, 2017, pp. 457–75, <https://doi.org/10.1111/capa.12246>.

Dunleavy, Patrick and others, 'The Theory of Modern Bureaucracy and the Neglected Role of IT', *Digital Era Governance: IT Corporations, the State, and e-Government* (Oxford, 2006; online edition, Oxford Academic, 3 Oct. 2011)

Engstrom, David Freeman and Ho, Daniel E., Artificially Intelligent Government: A Review and Agenda (March 9, 2020). *Big Data Law* (Roland Vogl, ed., 2020), Available at SSRN: <https://ssrn.com/abstract=3551549>

Fatima, Samar, et al. "National Strategic Artificial Intelligence Plans: A Multi-Dimensional Analysis." *Economic Analysis and Policy*, vol. 67, 2020, pp. 178–94, <https://doi.org/10.1016/j.eap.2020.07.008>.

Haque, Akhlaque. *Surveillance, Transparency, and Democracy : Public Administration in the Information Age*. The University Alabama Press, 2015.

Hildebrandt, Mireille. “The Issue Of Bias: The Framing Powers Of Machine Learning” *Machines We Trust : Perspectives on Dependable AI*. Edited by Marcello Pelillo and Teresa Scantamburlo, The MIT Press, 2021.

ISED (Innovation Science and Economic Development), Advisory Council on A. I. (2021, July). *Advisory Council on AI Annual report 2020–21*. Government of Canada, Innovation, Science and Economic Development Canada, Minister of Innovation, Science and Economic Development. Retrieved January 29, 2023, from <https://ised-isde.canada.ca/site/advisory-council-artificial-intelligence/en/annual-reports/annual-report-2020-21>

ISED (Innovation Science and Economic Development), Advisory Council on A. I. (2022, November) “Meeting Summaries: Advisory Council on Artificial Intelligence.” *Innovation, Science and Economic Development Canada*. Retrieved January 29, 2023, from <https://ised-isde.canada.ca/site/advisory-council-artificial-intelligence/en/meeting-summaries>

Leblanc, D. (2023, January 30). *Ottawa undermined Canadians' language rights with ArriveCan app, Language commissioner says | CBC News*. CBCnews. Retrieved January 30, 2023, from <https://www.cbc.ca/news/canada/montreal/arrivecan-app-language-bugs-1.6729130>

Longo, Justin. “When Artificial Intelligence Meets Real Public Administration.” *Canadian Public Administration*, vol. 65, no. 2, 2022, p. 384–, <https://doi.org/10.1111/capa.1465>.

Newell, Sue, and Marco Marabelli. “Strategic Opportunities (and Challenges) of Algorithmic Decision-Making: A Call for Action on the Long-Term Societal Effects of ‘datification.’”

The Journal of Strategic Information Systems, vol. 24, no. 1, 2015, pp. 3–14,  
<https://doi.org/10.1016/j.jsis.2015.02.001>.

Office of the Auditor General of Canada (2018, March). *Report 1 - Building and implementing the Phoenix Pay System*. Government of Canada, Office of the Auditor General of Canada. Retrieved January 30, 2023, from [https://www.oag-bvg.gc.ca/internet/English/att\\_\\_e\\_43045.html](https://www.oag-bvg.gc.ca/internet/English/att__e_43045.html)

PCO (Privy Council Office) (2020). Government of Canada. Canada.ca. Retrieved January 29, 2023, from <https://www.canada.ca/en/privy-council/services/blueprint-2020/beyond-2020.html>

Raji, I. D. (2022, July). *A culture of ethical AI: Report - CIFAR*. Retrieved January 30, 2023, from <https://cifar.ca/wp-content/uploads/2022/08/CIFAR-AI-Insights-EN-FINAL-AODA.pdf>

TBS (Treasury Board Secretariat) (2019, April). *Directive On Automated Decision-Making*. Canada.ca. Retrieved January 29, 2023, from <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>

“Third Review Of The Directive On Automated Decision-Making.” GCWiki. (2022). Retrieved January 29, 2023, from [https://wiki.gccollab.ca/Third\\_Review\\_of\\_the\\_Directive\\_on\\_Automated\\_Decision-Making](https://wiki.gccollab.ca/Third_Review_of_the_Directive_on_Automated_Decision-Making)  
 Valle-Cruz, David, et al. “Assessing the Public Policy-Cycle Framework in the Age of Artificial Intelligence: From Agenda-Setting to Policy Evaluation.” Government Information Quarterly, vol. 37, no. 4, 2020, p. 101509–,  
<https://doi.org/10.1016/j.giq.2020.101509>.

Wirtz, Bernd W., et al. “The Dark Sides of Artificial Intelligence: An Integrated AI Governance Framework for Public Administration.” *International Journal of Public Administration*, vol. 43, no. 9, 2020, pp. 818–29, <https://doi.org/10.1080/01900692.2020.1749851>.

Yampolskiy, Roman V. “Unexplainability and Incomprehensibility of AI.” *Journal of Artificial Intelligence and Consciousness*, vol. 07, no. 02, 2020, pp. 277–291.,  
doi:10.1142/s2705078520500150.