
September 22, 2022

Blair Attard-Frost
PhD Candidate
University of Toronto, Faculty of Information
blair@blairaf.com

Background
I was recently invited by staff of the Treasury Board Secretariat of Canada (TBS) to submit my comments on their 3rd Review of the Directive on Automated Decision-Making (DADM)\(^1\) and its associated Algorithmic Impact Assessment (AIA) tool.\(^2\) I was glad to be invited to provide my feedback on the 3rd Review, as the DADM is an important instrument for governing automated decision systems (ADS) in Canada’s public service and serves as a valuable model for other jurisdictions and organizations. I was invited to review: a slide deck summarizing key issues, policy recommendations, and provisional


amendments; an overview of relevant issues and changes; a report on the 3rd Review; and a consultations page on the Government of Canada’s GCwiki. I was invited by TBS staff to review those materials on September 12, 2022. I was told to submit my comments on those materials by no later than September 23, 2022. I submitted my comments to ai-ia@tbs-sct.gc.ca on September 22, 2022. I have decided to publish the comments that I submitted to TBS to ensure that my comments can serve as a public resource to any researchers, public servants, or other individuals or organizations who might benefit from reading my comments or from using them in their own work.

Preamble

The amendments that TBS has proposed to the DADM and AIA as part of the 3rd Review are highly commendable. The amendments will do much to improve the clarity, transparency, and overall effectiveness of the DADM and AIA. These improvements are vital to TBS's pursuits of responsible AI and bolstering public trust in the DADM. However, there are many additional opportunities beyond those identified in the 3rd Review for TBS to further improve the clarity, transparency, precision, and overall effectiveness of the instrument's application, and thereby bolster public trust in the DADM, in its application, and in its outcomes.

Below are the comments I have prepared in response to the 11 issues and proposed changes described within the 3rd Review of the DADM. Additionally, I have prepared some comments pertaining to an additional 2 issues (“Implementation of Quality Assurance” and “Public Access & Engagement”) that I believe the 3rd Review does not account for in sufficient detail.

My comments on those issues touch upon 20 key topics: (1) Scope of services, (2) Scope of agencies, (3) Scope of National Security Systems, (4) Scope of system lifecycle phases, (5) Review timeframe, (6)

---


Scope

- **Scope of services:** I strongly support expanding the scope of the Directive to also cover the internal services of federal institutions.

- **Scope of agencies:** To improve the overall clarity and transparency of the DADM—as well as to bolster public trust in the initiative—I support the suggestion made by Ana Brandusescu and Renée Sieber that the DADM should require a publicly accessible list or index of scheduled departments and agencies to be published and regularly updated.\(^7\) The list should clearly indicate which departments and agencies (1) fall under the scope of the DADM as per Section 9.1, (2) do not fall under the scope of or are excluded from the DADM as per Sections 8.2 and 9.1.1, and (3) have entered into a Specific Agreement with TBS as per Section 9.2, as well as the terms of each of those Specific Agreements.

- **Scope of National Security Systems:** The definition and scope of "National Security Systems" as described in Section 5.4 of the DADM is unclear. The DADM references the Policy on Service and Digital\(^8\) on this point, which uses the phrase "national security systems" (non-capitalized) in Section 5.6.3 and "National Security Systems" (capitalized) in Section 6.3, but this instrument does not describe a clear definition or scope of these types of systems either. Specifically, it is unclear (especially given the variations in capitalization) whether this term refers to: (a) some of the systems owned by departments or agencies with a clear and significant national security

---


mandate, such as the RCMP, DND, or CSIS; (b) all of the systems owned by departments or agencies with a clear and significant national security mandate; or (c) potentially any system throughout the federal government that is determined to require special protections or exemptions from the DADM in the interest of national security (however that might be determined/defined). In the absence of a clear, public-facing definition of "National Security Systems" as specified by another instrument, TBS should take initiative to clearly indicate in the DADM the criteria that an ADS must meet in order to qualify as a “National Security System.” This could be added either to Section 5.4 or as a new definition to Appendix A.

- **Scope of system lifecycle phases:** The DADM should clearly define the terms it uses to identify and differentiate the various phases of the lifecycle of an ADS. At present, it is not entirely clear what the typical activities and goals involved in many of the phases described within the DADM are, as well as how the various phases relate to one another and to the provisions of the DADM. To improve the clarity of the DADM’s scope and to encourage transparency, consistency, and precision in the DADM’s application, the specific operational activities and goals involved in the following phases should be more clearly defined:

  - "Use" as described in Sections 4.2.3, 6.3.8, 8.1, and 8.4.
  - "Production" as described in Sections 5.3, 6.1.1, 6.3.1, and 6.3.10.
  - "Design" and "implementation" as described in Section 6.3.5.
  - "Development cycle" as described in Section 6.3.7.
  - "Concept stage” as described in Section 6.3.8.
  - The pre-production phase(s) referenced in Sections 6.1.1 and 6.3.10 ("Prior to the production") as well as in Section 6.3.1 ("Before launching into production"). Reconciling the scope of pre-production with the scope of production is particularly important, as Section 5.3 seems to suggest that as activities that exist outside of production (in the language of the DADM, "before" and "prior to" production), all pre-production activities
are therefore beyond the scope of the DADM, which is inconsistent with the provisions of Sections 6.1.1, 6.3.1, and 6.3.10.

- "Test"/"Testing" as described in Sections 6.2.5.2, 6.3, and 6.3.1 as well as "test environments" as described in Section 5.3. The definition of "test environment" provided in Appendix A is helpful for communicating what the DADM views to be the activities and goals involved in the operation of a typical test environment, but the use of this term in Section 5.3 needs to be reconciled with the use of "test"/"testing" elsewhere. Section 5.3 seems to suggest that if at any point an ADS operates within any type of test environment, it will then be operating entirely outside the scope of DADM, which is inconsistent with the provisions of Sections 5.3, 6.2.5.2, 6.3, and 6.3.1. Greater specificity about the conditions under which an ADS in a test environment might be excluded from the DADM as per Section 5.3 and how such exclusion would not conflict with the requirements the DADM sets regarding testing throughout Sections 6.2 and 6.3 would improve the DADM's clarity and transparency.

- These terms could be clarified within the relevant sections of the DADM, as new definitions added to Appendix A, and/or by including a diagram of the ADS lifecycle as it is understood in the language of the DADM in a new appendix.

**Periodic Review**

- **Review timeframe**: I agree with expanding the timeframe of the periodic review to allow more time for TBS staff to conduct evidence-gathering, compliance assessment, and amendment processes. However, I share the sentiment of Ana Brandusescu and Renée Sieber that a 2 year timeframe is too long given the rapid pace at which new research on artificial intelligence is produced, new AI systems are produced, the capabilities of automated decision systems change, and the potential social and environmental impacts of automated decision-making increase in scale. To better contend with that rapid pace, TBS ought to conduct a periodic review of the DADM every 1 year for the time being and, if necessary, allocate additional resources to DADM-related initiatives and staff to support with DADM monitoring, reporting, and improvement activities. The review timeframe should then be scaled back to more than 1 year in the future.
once more opportunities to improve the DADM have been implemented, DADM-related operations become more routine within TBS and other federal institutions, and the Government of Canada is better operationally prepared to contend with the rapid pace of AI production and AI impact.

Clients Impacted

- **Human-centred language**: I agree with TBS's rationale for replacing "Canadians" with a more broadly representative term, but I disagree with proposed use of the term "clients." Applying a transactional, “business-minded” language to human beings and our political relations has significant practical utility, so it has understandably become standard practice in the realm of public service.\(^9\) However, this language evokes a dehumanized vision of public service in which the relationship between human beings and our democratic institutions is reduced to calculative client-provider or client-server transactions. Re-envisioning public service as little more than a set of business transactions does little to bolster public trust, and at worst, risks undermining public trust (particularly when directed at the public through public-facing documentation such as the DADM). While I do prefer the term “Canadian society” over “clients,” the former term can unfortunately be interpreted as excluding people who do not reside in Canada (e.g., visa applicants and other people living abroad who are seeking service from the federal government). Therefore, in Section 4.1, I recommend replacing "Canadians and federal institutions" with "federal institutions and people seeking service from the Government of Canada." More generally, I recommend replacing "Canadians" with "People seeking service from the Government of Canada", or "people seeking service from a federal institution", or "people to whom the Government of Canada is responsible."

Data Governance

- I agree with the proposed addition of a requirement to govern the data used and generated by automated decision systems. Good data governance is an essential component of good AI governance, so the addition of a new sub-section to Section 6.3 to set specific quality assurance

---

requirements around data privacy, security, retention, and disposition is vital to the
effectiveness of the DADM.

Model Bias

- **Human & institutional biases in application**: I share the sentiment of Ana Branduseascu and Renée Sieber that testing the inherent bias of models is an important addition to the DADM, but debiasing data or models does not necessarily offset an ADS’s capacity to disproportionately distribute harms *within the context of its expected application*. In addition to data bias and model bias, TBS should also recognize that human and institutional biases can be inherent in the expected context of the system’s application.¹⁰ For example, the application of facial recognition and other automated tools to policing is an extremely bias-prone practice, often serving to reinforce the institutional biases and injustices that have historically been perpetuated against disadvantaged peoples and communities.¹¹ As another example, automated tools used to make employment decisions are also known to be extremely bias-prone due to historic human and institutional biases in employment decision-making.¹² The European Union’s *Artificial Intelligence Act*¹³ recognizes these principles of inherent bias and inherent risk by stratifying system risk into multiple categories that are based on the expected context of the system’s practical application. I strongly suggest that both the DADM and the AIA be amended to recognize that some categories of ADS applications carry inherent risks of human and institutional bias that cannot be accurately determined or corrected through an assessment of risk areas 1-6 or mitigation areas 1-2 as they currently exist in the AIA. TBS ought to undertake a thorough review of other policy instruments and research within Canada and internationally.

---


that (a) provide an inventory of ADS applications that have inherently high risk profiles due to their tendency to promote and reinforce human and institutional biases, and (b) outline application-sensitive approaches to AI risk assessment and mitigation. This will enable TBS to build an application-sensitive approach into the DADM, as well as new risk areas and mitigation areas into the AIA that are specifically designed to counteract human and institutional bias in the expected context of the system’s application. Such an approach might include, for example:

- The addition of provisions to Appendix C of the DADM and mitigation area 1 of the AIA to explicitly require a greater range of affected non-expert stakeholders (such as people to whom the ADS is expected to provide service, people to whom the ADS is expected to pose risks, relevant data subjects, and the general public) to be consulted and included in the design and evaluation of certain high-risk systems, as called for by participatory design methodologies such as participatory AI, value-sensitive design, and design justice.

- For suppliers involved in the design and production of high-risk systems: assessments of the supplier’s values statements, organizational culture, business and legal history, record of conformance with international human rights law, and record of conformance with the Government of Canada’s guiding principles on effective and ethical use of AI.

- For suppliers involved in the design and production of high-risk systems: review of records of initiative outcomes for any training or change management initiatives related to diversity, equity, inclusion, bias, and/or ethics.

---


The addition of new provisions to Section 6.3.5 that better specify the definition of “adequate” employee training, and to require that employee training specifically on the ethics of ADS design, function, and implementation is successfully provided prior to the system being used.

- These changes will ultimately improve the effectiveness of both instruments in mitigating bias at the levels of data, model, and application, as well as the efficiency of TBS in categorizing, assessing, and managing ADS risk.

Inclusion

- **Intersectionality**: I support the addition of an intersectional analysis methodology to the DADM, but TBS should recognize that concerns have been expressed that the proposed GBA+ methodology “instigates fundamental misunderstandings of intersectionality” and functions as an “expert-bureaucratic model” of gender-based analysis rather than as a “participatory-democratic” model. TBS should therefore review the GBA+ methodology and compare its strengths and weaknesses with those of participatory methodologies for intersectional design and analysis.

- **Disproportionality of environmental harms**: I also suggest that if TBS wishes to pursue an intersectional approach to ADS design, development, use, then TBS ought to consider the capacity of ADS and other data-intensive AI systems (particularly those that make use of natural language processing) to cause negative environmental impacts, primarily due to their intensive energy usage during operation. An account of environmental impact is especially important for

---


an intersectional approach, as environmental harms such as climate change and water depletion are well-known to disproportionately impact historically disadvantaged peoples and communities. Fortunately, many potential solutions for mitigating the environmental risks posed by data-intensive AI systems have been proposed.\textsuperscript{21} TBS ought to conduct a thorough review of such solutions so that they can be built into future versions of the DADM as well as into the risk areas and mitigation areas of the AIA.

- \textit{Inclusive design}: I strongly suggest that TBS include a broader range of stakeholders who may potentially be impacted by ADS production in the design and quality assurance activities of all relevant systems, and particularly in the design and quality assurance of high-risk systems. These stakeholders may include, for example, people to whom the ADS is expected to provide service, people to whom the ADS is expected to pose risks, relevant data subjects, or other members of the general public who do not necessarily have AI or policy expertise but have a personal stake in the system nonetheless. Participatory design methodologies such as those described in my above comments on “Model Bias” should be adopted by TBS to ensure that disadvantaged peoples and communities who are at risk of being harmed by the ADS are included in design and evaluation activities that are relevant to their stake in their system.

\textbf{Explanation}

- \textit{Explainability by design}: I agree with the proposed amendments that specify in greater detail the criteria of a meaningful explanation. I also echo the comment made by Ana Brandusescu and Renée Sieber that ideally, explainability would be considered and built into the ADS during the design phase. Additionally, I will add that specifying the phases and activities involved in the system lifecycle in greater detail (as I suggest in my above comments on "Scope") will enable the DADM to set clearer and more precise requirements targeted at specific phases and activities in the system lifecycle (e.g., clearer and more precise explainability requirements that apply specifically to the design phase vs. explainability requirements that apply specifically to the use phase).

Reasons for Automation

- I support the proposed amendments to the AIA regarding reasons for automation. Ensuring that ADS in federal institutions provide a clear public benefit, are responsive to user needs, and have been compared with alternative non-automated solutions is an important function for both the DADM and the AIA.

Peer Review

- **Public ADS registry:** I strongly agree with the proposed amendments to mandate publication of peer reviews and completion of peer reviews prior to system production. I also agree with the suggestion of Ana Brandusescu and Renée Sieber to create a publicly accessible ADS registry. Ideally, such a registry will contain an inventory of all ADS in the federal government that fall within the scope of DADM, and for each ADS, provide an entry containing the following:

  - A general description of the system’s intended function, intended user(s), designer(s)/developer(s), supplier(s), and date launched into production.

  - Any AIAAs associated with the system.

  - Any peer reviews associated with the system.

  - Procurement information and records associated with the system (RFPs, proposals, vendor communications, etc.).

  - Information and records pertaining to decision explanations, quality assurance, recourse outcomes, and system performance that was produced in support of the monitoring and reporting requirements of Sections 6.2.8, 6.3, 6.4.1, and 6.5.1.

  - Building upon the suggestions related to AIA findability and usability made by Teresa Scassa,²² the registry should also offer users the ability to filter and sort registered ADS.

---

systems according to various criteria such as risk/impact score, department or agency, supplier, or date launched into production.

Contingency

- I agree that re-framing Section 6.3.6 in terms of IT and business continuity management is a sensible amendment to improve the clarity of the section.

Timing of AIA Release

- I support the proposed amendment to complete and release the final results of an AIA prior to system production.

Implementation of Quality Assurance

- I strongly agree with the suggestion of Ana Brandusescu and Renée Sieber that the DADM should “include more specific, ongoing monitoring and reporting requirements so the public knows if the use or impact of an AI system has changed since the initial AIA.”

- **Transparency of QA activities:** At present, it is not clear if the quality assurance activities described in Section 6.3 of the DADM as well as the explanation activities described in Section 6.2.3 are being routinely implemented in practice. Specifically, the current state of the following activities is unclear:

  - The explanation requirement described in Section 6.2.3 specifies that a meaningful explanation of a decision will be provided to affected individuals, but it is not clear if any relevant decisions have been made to date, and if so, whether meaningful explanations are being routinely provided to affected individuals.

The monitoring requirement described in Section 6.3.2 specifies that ADS outcomes will be monitored and their compliance with applicable legislation verified on a “scheduled basis,” but it is not clear what that monitoring schedule is, whether it is currently being observed and implemented in practice, and—if it is being implemented in practice—what the results of routine ADS monitoring have been to date.

The data quality validation requirement described in Section 6.3.3 specifies that validation must be performed to ensure that the data collected for and used an ADS is relevant and up-to-date, but it is not clear if those validation activities have been performed in the past, how often they are being repeated, or whether that data remains relevant and up-to-date today.

The employee training requirement described in Section 6.3.5 specifies that employee training in ADS design, function, and implementation must be provided, but it is unclear if any such training has actually been provided, and if so, what form that training has taken and what the outcomes of the training have been.

The risk assessment requirement described in Section 6.3.7 specifies that multiple risk assessments must be conducted during the development cycle, but it is unclear if any of these risk assessments are being conducted, and if so, what form they have taken and what their findings have been.

The legal consultation requirement described in Section 6.3.8 specifies that consultation with an institution’s legal services must be undertaken “from the concept stage of a project,” but it is unclear whether any of these required legal consultations have been or are being conducted, and if so, what the outcomes of the consultations have been.

The recourse requirement described in Section 6.4.1 specifies that recourse options are available to challenge decisions, but it is unclear if any individuals or institutions have challenged any decisions to date. In addition, it is unclear what types of recourse options TBS believes may be “applicable” or inapplicable given the decisions that they expect in-scope systems to be making.
- The reporting requirement described in Section 6.5.1 specifies that information on the effectiveness and efficiency of every in-scope ADS in meeting program objectives must be published, but to date, no such information has been published. In addition, there is not a clear schedule specifying how regularly such information should be reviewed and published, no clear requirements specifying the structure and procedures the reporting process must observe, as well as no clear requirements specifying the types of findings that must be reported on.

- **QA of expected vs. actual use:** A consequence of the lack of transparency regarding whether or not the above quality assurance practices are being routinely implemented is that, from the perspective of a member of the public with access to only public-facing documentation, the DADM presently seems to be functioning as an instrument to only predict quality based on the expected use of in-scope systems prior to their production rather than to assure quality during the actual use of in-production systems. In other words, from a public perspective, the DADM does not currently appear to be fully upholding its commitments to quality assurance and risk mitigation in its application and in its outcomes.

- **QA scheduling & reporting requirements:** In the interest of bolstering public trust in the application and outcomes of the DADM, I suggest that TBS produce more clear and detailed monitoring and reporting requirements that include routine QA activity schedules for all in-scope ADS. At minimum, such requirements should mandate departments and agencies to regularly publish reports describing in detail the efforts that they have made to comply with the QA requirements. Ideally, such requirements should also mandate departments and agencies to regularly publish QA-related records that are deemed “appropriate” for disclosure according to the appropriateness criteria specified in an amended version of Section 4.2.3 (see my below comments on “Public Access & Engagement”). The details and schedules of these new monitoring and reporting requirements should be amended into the DADM as new sub-sections within Section 6, particularly within Section 6.3. Routine, scalable, and robust quality assurance practices are crucial to the success of the DADM, and will only become all the more crucial as ADS capabilities continue to evolve, costs associated with ADS development and use continue to drop, and more systems continue to be put in production within federal institutions.
• Resources for QA: TBS should allocate more resources to the DADM and its associated initiatives and staff in order to ensure that quality assurance activities can be routinely implemented at scale, effectively reported on, and published in a transparent and timely manner.

Public Access & Engagement

• Standards of public access and disclosure: The above comments I have made regarding limitations in public access to QA-related information and records point to a more general issue: at present, TBS is not providing a reasonable standard of public access to information and records related to the DADM’s application and outcomes. I strongly suggest that in addition to implementing the solutions to public access limitations that I have proposed throughout my comments, TBS should also build a more specific set of provisions and requirements into the DADM that more precisely specify the criteria used to determine when public disclosure of ADS-related data and information is “appropriate” as per Section 4.2.3.

• Fairness, accountability, and transparency in public engagement: Limitations in procedural fairness, public accountability, and transparency in current DADM-related stakeholder engagement processes—and particularly in the engagement activities involved in the DADM’s periodic reviews—pose an additional challenge for TBS in securing greater public trust in the DADM, its application, and its outcomes. I commend TBS for their hard work in undertaking a lengthy stakeholder engagement process as part of the 3rd Review, as well as for publishing a summary of their Phase 1 engagement findings on the GCwiki. However, there remain multiple opportunities for improving how future engagements are conducted and for thereby fostering greater public trust in the DADM, its application, and its outcomes. For example:
  
  o In order to bolster public trust, consultations on amending the DADM ought to be open to the general public, and the general public ought to be encouraged to submit comments and feedback to TBS. As a component of the public consultation process, TBS could also continue to conduct targeted engagements or send targeted participation invitations to specific researchers, members of civil society, other government agencies,
or other organizations whose comments TBS would in particular like to consider as part of their review. The Office of the Privacy Commissioner of Canada as well as the Ontario Digital Service are two notable examples of Canadian agencies that have recently conducted effective public consultations on digital, data, and AI policy issues.

- Following the example of the Ontario Digital Service’s public consultation on *Artificial Intelligence (AI) Guidance*, TBS should also conduct roundtable or town hall style public consultations. Feedback obtained from public consultations should then compiled, published, and encouraged on a continuous basis.

- A fair and reasonable timeframe should be given for all members of the public to review and reflect upon the relevant documentation, to write their comments, and to submit their comments. Alloting some members of the public 2 months to submit their comments on the 3rd Review while allotting other members of the public a significantly shorter timeframe to submit their comments undermines the DADM’s commitment to principles of “transparency, accountability,” and “procedural fairness.” TBS should ensure that all members of the public have an equal and reasonable timeframe to reflect upon the relevant documentation and submit their comments.

- In the interest of bolstering public trust in the DADM itself, public consultation and engagement activities should not be treated as peripheral to DADM policy. Instead, a precise set and schedule of public engagement activities that are required to be performed as part of every periodic review should be added directly into the DADM (e.g., into Section 1.2 or 6.3). This will enable TBS to better assure quality and public accountability in the review process.

---


• **Review of public consultation and engagement opportunities:** I suggest that TBS conduct a thorough review of these and other opportunities to improve future public consultation and engagement processes. Ideally, such a review would itself involve a great deal of public consultation and engagement in order to determine how best to ensure that the public’s voice is better accounted for in the DADM, in its application, and in its outcomes.

• **Resources for public consultation and engagement:** TBS should allocate more resources to the DADM and its associated initiatives and staff in order to ensure that public consultation and engagement activities can be routinely implemented at scale, effectively reported on, and published in a transparent and timely manner.