Algorithmic Impact Assessment Results

Version: 0.10.0

Project Details

1. Name of Respondent
Hamza Chikhaoui

2. Job Title
Prioritize Co-founder

3. Department
Health (Department of)

4. Branch
N/A

5. Project Title
TriageAssist

6. Project ID from IT Plan
N/A

7. Departmental Program (from Department Results Framework)
N/A

8. Project Phase
Implementation [Points: 0]

9. Please provide a project description:
We have decided during the COVID-19 epidemic to create an application that allows nurses to triage patients within a resource constrained hospital environment.

The triage is made via a machine learning algorithm that classifies patients based on whether they are likely or not to have a heart disease. The main objective of this classifier is to help medical professionals to decide whether patients should be prioritized in terms of the care they receive or not.

Nurses and doctors can in turn use the classifier to see whether a new patient at the ER is likely to have a heart disease or not, and use this information to make informed decisions on the treatment of the patient. Indeed, if the algorithm classifies someone as having a heart disease, then the ER can take more time with their diagnosis and treatment decision. If not, they can speed up the process.

About The System

10. Please check which of the following capabilities apply to your system.
Risk assessment: Analyzing very large data sets to identify patterns and recommend courses of action and in some cases trigger specific actions
Content generation: Analyzing large data sets to categorize, process, triage,

personalize, and serve specific content for specific contexts

Section 1: Impact Level : 2

Current Score: 47

Raw Impact Score: 55

Mitigation Score: 40

Section 2: Requirements Specific to Impact Level 2

Peer review

Consult at least one of the following experts and publish the complete review or a plain language summary of the findings on a Government of Canada website:

- qualified expert from a federal, provincial, territorial or municipal government institution
- qualified members of faculty of a post-secondary institution
- qualified researchers from a relevant non-governmental organization
- contracted third-party vendor with a relevant specialization
- a data and automation advisory board specified by Treasury Board of Canada Secretariat.

OR

Publish specifications of the automated decision system in a peer-reviewed journal. Where access to the published review is restricted, ensure that a plain language summary of the findings is openly available.

Gender-based Analysis Plus

Ensure that the Gender-based Analysis Plus addresses the following issues:

- impacts of the automation project (including the system, data and decision) on gender and/or other identity factors;
- planned or existing measures to address risks identified through the Gender-based Analysis Plus.

Notice

Plain language notice posted through all service delivery channels in use (Internet, in person, mail or telephone).

Human-in-the-loop for decisions

Decisions may be rendered without direct human involvement.

Explanation

In addition to any applicable legal requirement, ensure that a meaningful explanation is provided to the client with any decision that results in the denial of a benefit or service, or involves a regulatory action. The explanation must inform the client in plain language of:

- the role of the system in the decision-making process;
- the training and client data, their source, and method of collection, as applicable;
- the criteria used to evaluate client data and the operations applied to process it;
- the output produced by the system and any relevant information needed to interpret it in the context of the administrative decision; and
- a justification of the administrative decision, including the principal factors that led to it.

Explanations must also inform clients of relevant recourse options, where appropriate.

A general description of these elements must also be made available through the Algorithmic Impact Assessment and discoverable via a departmental website.

Training

Documentation on the design and functionality of the system.

IT and business continuity management

None

Approval for the system to operate

None

Other requirements

The Directive on Automated Decision-Making also includes other requirements that must be met for all impact levels.

Link to the Directive on Automated Decision-Making

Contact your institution's ATIP office to discuss the requirement for a Privacy Impact Assessment as per the Directive on Privacy Impact Assessment.

Section 3: Questions and Answers

Section 3.1: Impact Questions and Answers

Reasons for Automation

1. What is motivating your team to introduce automation into this decision-making process? (Check all that apply)
Improve overall quality of decisions
Use innovative approaches
Existing backlog of work or cases
Lower transaction costs of an existing program

2. What client needs will the system address and how will this system meet them? If possible, describe how client needs have been identified.
The clients needs that the system address are the following:
- The need to receive appropriate care when getting to the ER depending or whether or not they have a heart disease.

- The need to be receiving care as fast as possible, as the algorithm greatly speeds up the triaging process at the hospital.
- The need to be receiving quality care, as more time would be spent by the healthcare practitioners actually healing the patients rather than undertaking logistics tasks

3. Please describe any public benefits the system is expected to have.
The public benefits of the system would be that they can alleviate some critical pressure on hospitals by improving the flow of patient management , allow for a faster and more tailored care of the patients, and improve the experience of patients when receiving urgent care.

4. How effective will the system likely be in meeting client needs?
Very effective [Points: +0]

5. Please describe any improvements, benefits, or advantages you expect from using an automated system. This could include relevant program indicators and performance targets.
The following improvements are to be noted through our use of an automated system:
- A more systematic and more reliable triaging system.
- A faster triaging system that human triaging.
- Allowing more patients to be taken care of through prioritizing the patients require more urgent care

6. Please describe how you will ensure that the system is confined to addressing the client needs identified above.
The system is confined to addressing the clients needs as the application comports only the classifier as a core module. The algorithm's sole purpose is to accurately classify the patients that are prone to having a heart disease. Furthermore, the algorithm is to be used under the supervision of a health practitioner that will oversee the triaging process.

7. Please describe any trade-offs between client interests and program objectives that you have considered during the design of the project.
There is a tradeoff between every patient wanting to receiving urgent care when getting at the ER, and the need for our model (and hospitals) to prioritize that patients that are prone to having a heart disease.

8. Have alternative non-automated processes been considered?
No [Points: +1]

9. What would be the consequence of not deploying the system?
Service costs are too high [Points: 0]
Service quality is not as high [Points: 0]
Service cannot be delivered in a timely or efficient manner [Points: +2]

Risk Profile

10. Is the project within an area of intense public scrutiny (e.g. because of privacy concerns) and/or frequent litigation?
Yes [Points: +3]

11. Are clients in this line of business particularly vulnerable?

Yes [Points: +3]

12. Are stakes of the decisions very high?
Yes [Points: +4]

13. Will this project have major impacts on staff, either in terms of their numbers or their roles?
Yes [Points: +3]

14. Will the use of the system create or exacerbate barriers for persons with disabilities?
No [Points: +0]

Project Authority

15. Will you require new policy authority for this project?
No [Points: +0]

About the Algorithm

16. The algorithm used will be a (trade) secret
Yes [Points: +3]

17. The algorithmic process will be difficult to interpret or to explain
No [Points: +0]

About the Decision

18. Please describe the decision(s) that will be automated.
The decision that will be automated is the decision of whether or not a patient should receive urgent care at the ER. This decision is the consequence of the patient being classified or not as likely to have a heart disease by the classifier.

19. Does the decision pertain to any of the categories below (check all that apply):
Health related services [Points: +1]

Impact Assessment

20. Which of the following best describes the type of automation you are planning?
Partial automation (the system will contribute to administrative decision-making by supporting an officer through assessments, recommendations, intermediate decisions, or other outputs) [Points: +2]

21. Please describe the role of the system in the decision-making process.
The system will be used to predict to the health practitioner whether or not the patient is likely to have a heart disease or not. However, the final decision of whether or not the patient will receive urgent care at the ER will be made by the health practitionner.

22. Will the system be making decisions or assessments that require judgement or discretion?
No [Points: +0]

23. Please describe the criteria used to evaluate client data and the operations applied to process it.

The algorithm will use a machine learning model to classify whether or not the patient is at risk of a heart disease.

24. Please describe the output produced by the system and any relevant information needed to interpret it in the context of the administrative decision.
The output of the machine learning algorithm is the positive prediction (1) or not (0) of a heart disease in a patient according to a series of features. No other information is needed to interpret the results of the algorithm

25. Will the system perform an assessment or other operation that would not otherwise be completed by a human?
No [Points: +0]

26. Is the system used by a different part of the organization than the ones who developed it?
Yes [Points: +4]

27. Are the impacts resulting from the decision reversible?
Reversible [Points: +1]

28. How long will impacts from the decision last?
Some impacts may last a matter of months, but some lingering impacts may last longer [Points: +2]

29. Please describe why the impacts resulting from the decision are as per selected option above.
We can start off by saying that the impacts from the decision are reversible as the health practitioners can at any moment change their decisions and provide a patient with urgent care.

Moreover, the timing of the impacts of the decision are to be analyzed on a case by case basis, as a wrong decision from the algorithm can have immediate impacts that are unrelated to the algorithm (for example the state of a patient can deteriorate and require urgent care that would be unrelated a heart disease, or the algorithm). On the other hand, correct decisions made by the algorithm can potentially save a patient's life, but do not constitute a guarantee of him being well in the coming months or years.

Therefore, the timeline of the impacts is the product of other events and cannot be solely quantified as a result of the correct/incorrect decision made by the algorithm.

30. The impacts that the decision will have on the rights or freedoms of individuals will likely be:
Little to no impact [Points: +1]

31. Please describe why the impacts resulting from the decision are as per selected option above.
The freedoms of individuals are not affected by the algorithm, as our model does not have any consequence on the freedom of patients to access healthcare. Moreover, the patient can decide at any point to opt-out and receive care from different practitioners.

One could nuance this statement. The output of the algorithm does not constitute a barrier to access healthcare services for the patient per say.

However, the algorithm may affect the speed at which the patient receives care (whether faster or "slower") which in situations of urgency may be critical.

32. The impacts that the decision will have on the equality, dignity, privacy, and autonomy of individuals will likely be:
Little to no impact [Points: +1]

33. Please describe why the impacts resulting from the decision are as per selected option above.
There is inherent bias in any algorithm that would be based on existing data. However, we have put in place a series of measures to limit the bias of our algorithm with regards to the different populations represented in our dataset. We have also put in place privacy measures to ensure that protecting the privacy of the patients is at the core of our model

34. The impacts that the decision will have on the health and well-being of individuals will likely be:
High impact [Points: +3]

35. Please describe why the impacts resulting from the decision are as per selected option above.
A correct decision from the algorithm can have as a consequence to save the life of a patient. On the other hand, a wrong decision made by the algorithm (not detecting a heart disease for example) may in turn cause dramatic consequences for the patient.
Because this technology is designed to be implemented in hospitals, the hopefully positive impacts of the algorithm will have a direct impact on the health and well-being of the patients, by allowing health practitioners to have more time to provide critical care to the patients.

36. The impacts that the decision will have on the economic interests of individuals will likely be:
Little to no impact [Points: +1]

37. Please describe why the impacts resulting from the decision are as per selected option above.
Because our model is to be implemented under the direct supervision of healthcare practitioners, one can hardly see how the model will replace some health practitioners' jobs or cause layoffs from the hospital.
Moreover, there will be no additional direct impact on the economic interests of individuals other than the ones that already exist in a hospital setting

38. The impacts that the decision will have on the ongoing sustainability of an environmental ecosystem, will likely be:
Little to no impact [Points: +1]

39. Please describe why the impacts resulting from the decision are as per selected option above.
One can say that there is only little to no added energy consumption required to run our model. Moreover, one could argue that whatever little energy is used to implement the algorithm is widely counterbalanced by the savings made as a result of a better ressource allocation within the hospital.

About the Data - A. Data Source

40. Will the Automated Decision System use personal information as input data?
Yes [Points: +4]

41. Have you verified that the use of personal information is limited to only what is directly related to delivering a program or service?
Yes [Points: +0]

42. Is the personal information of individuals being used in a decision-making process that directly affects those individuals?
Yes [Points: +2]

43. Have you verified if the system is using personal information in a way that is consistent with: (a) the current Personal Information Banks (PIBs) and Privacy Impact Assessments (PIAs) of your programs or (b) planned or implemented modifications to the PIBs or PIAs that take new uses and processes into account?
No [Points: +1]

44. What is the highest security classification of the input data used by the system? (Select one)
Protected A [Points: +1]

45. Who controls the data?
Other Canadian Government (prov/municipal) [Points: +2]

46. Will the system use data from multiple different sources?
Yes [Points: +4]

47. Will the system require input data from an Internet- or telephony-connected device? (e.g. Internet of Things, sensor)
No [Points: +0]

48. Will the system interface with other IT systems?
No [Points: +0]

49. Who collected the data used for training the system?
A foreign government or non-government third party [Points: +4]

50. Who collected the input data used by the system?
Your institution [Points: +1]

51. Please describe the input data collected and used by the system, its source, and method of collection.
Health practitioners at the hospital where the system will be deployed will collect the input data from the patients for classification

About the Data - B. Type of Data

52. Will the system require the analysis of unstructured data to render a recommendation or a decision?
No [Points: 0]

Section 3.2: Mitigation Questions and Answers

Consultations

1. Internal Stakeholders (federal institutions, including the federal public service)
Yes [Points: +1]

2. Which Internal Stakeholders have you engaged?
Other (describe)

3. Please describe
We are engaged with the hospitals that will use our product, who themselves are subject to compliance on standards from healthcare governing entities on the provincial and federal level

4. External Stakeholders (groups in other sectors or jurisdictions)
Yes [Points: +1]

5. Which External Stakeholders have you engaged?
Civil Society
Academia

De-Risking and Mitigation Measures - Data Quality

6. Do you have documented processes in place to test datasets against biases and other unexpected outcomes? This could include experience in applying frameworks, methods, guidelines or other assessment tools.
Yes [Points: +2]

7. Is this information publicly available?
Yes [Points: +1]

8. Have you developed a process to document how data quality issues were resolved during the design process?
Yes [Points: +1]

9. Is this information publicly available?
Yes [Points: +1]

10. Have you undertaken a Gender Based Analysis Plus of the data?
Yes [Points: +1]

11. Is this information publicly available?
Yes [Points: +1]

12. Have you assigned accountability in your institution for the design, development, maintenance, and improvement of the system?
Yes [Points: +2]

13. Do you have a documented process to manage the risk that outdated or unreliable data is used to make an automated decision?
Yes [Points: +2]

14. Is this information publicly available?
Yes [Points: +1]

15. Is the data used for this system posted on the Open Government Portal?
No [Points: +0]

De-Risking and Mitigation Measures - Procedural Fairness

16. Does the audit trail identify the authority or delegated authority identified in legislation?
Yes [Points: +1]

17. Does the system provide an audit trail that records all the recommendations or decisions made by the system?
Yes [Points: +2]

18. Are all key decision points identifiable in the audit trail?
Yes [Points: +2]

19. Are all key decision points within the automated system's logic linked to the relevant legislation, policy or procedures?
Yes [Points: +1]

20. Do you maintain a current and up to date log detailing all of the changes made to the model and the system?
Yes [Points: +2]

21. Does the system's audit trail indicate all of the decision points made by the system?
Yes [Points: +1]

22. Can the audit trail generated by the system be used to help generate a notification of the decision (including a statement of reasons or other notifications) where required?
No [Points: +0]

23. Does the audit trail identify precisely which version of the system was used for each decision it supports?
Yes [Points: +2]

24. Does the audit trail show who an authorized decision-maker is?
Yes [Points: +1]

25. Is the system able to produce reasons for its decisions or recommendations when required?
No [Points: +0]

26. Is there a process in place to grant, monitor, and revoke access permission to the system?
Yes [Points: +1]

27. Is there a mechanism to capture feedback by users of the system?
Yes [Points: +1]

28. Is there a recourse process established for clients that wish to challenge the decision?
Yes [Points: +2]

29. Does the system enable human override of system decisions?
Yes [Points: +2]

30. Is there a process in place to log the instances when overrides were performed?
Yes [Points: +1]

31. Does the system's audit trail include change control processes to record modifications to the system's operation or performance?
Yes [Points: +2]

32. Have you prepared a concept case to the Government of Canada Enterprise Architecture Review Board?
Yes [Points: +1]

De-Risking and Mitigation Measures - Privacy

33. If your system uses or creates personal information, have you undertaken a Privacy Impact Assessment, or updated an existing one?
No [Points: +0]

34. Have you undertaken other types of privacy assessments for your automation project? Please describe any relevant efforts.
- We have implemented k anonymous and l diverse datasets.
- We have implemented differentially private models

35. Have you designed and built security and privacy into your systems from the concept stage of the project?
Yes [Points: +1]

36. Is the information used within a closed system (i.e. no connections to the Internet, Intranet or any other system)?
Yes [Points: +1]

37. If the sharing of personal information is involved, has an agreement or arrangement with appropriate safeguards been established?
Yes [Points: +1]

38. Will you de-identify any personal information used or created by the system at any point in the lifecycle?
Yes [Points: +1]

39. Please describe your de-identification method(s).
k anonimity, l diversity, differentially private models