WORKING PAPERS

Using Machine-Learning to Triage Canada's Temporary Resident Visa Applications

Lucia Nalbandian

Working Paper No. 2021/9
July 2021

The *Working Papers* Series is produced jointly by the Ryerson Centre for Immigration and Settlement (RCIS) and the CERC in Migration and Integration

www.ryerson.ca/rcis
www.ryerson.ca/cerc-migration

Ryerson University
Canada Excellence Research Chair in Migration & Integration
RCIS RYERSON CENTRE for Immigration and Settlement

Working Paper

No. 2021/9

Using Machine-Learning to Triage Canada's Temporary Resident Visa Applications

Lucia Nalbandian
Ryerson University

Series Editors: Anna Triandafyllidou and Usha George

The *Working Papers* Series is produced jointly by the Ryerson Centre for Immigration and Settlement (RCIS) and the CERC in Migration and Integration at Ryerson University.

Working Papers present scholarly research of all disciplines on issues related to immigration and settlement. The purpose is to stimulate discussion and collect feedback. The views expressed by the author(s) do not necessarily reflect those of the RCIS or the CERC.

For further information, visit www.ryerson.ca/rcis and www.ryerson.ca/cerc-migration.

ISSN: 1929-9915

Table of Contents

Introduction

New digital technologies have given rise to innovative ways to deliver services. Among these novel approaches, artificial intelligence has increasingly been viewed as a tool to be used in public administration and service delivery to efficiently implement public policies, roll out programs and deliver public services. The immigration sector is not immune to the allure of artificial intelligence and digital transformation. Rather, states are increasingly turning to advanced technologies for the management of migration. Accordingly, it is unsurprising that the Canadian government, through Immigration, Refugees and Citizenship Canada (IRCC) and the Immigration and Refugee Board, have recently begun exploring and introducing novel digital technologies to address demand and backlogs in Canada's immigration sector. As of 2020, IRCC has been operating three analytics-based systems to triage parts of intake of Temporary Resident Visa (TRV) applications, including a system for TRVs received online from China, a system for TRVs received online from India and finally a system for TRVs received from India through a Visa Application Centre. The two former systems were launched in 2018 and, together, form the focal point of this paper. Specifically, this paper establishes how the Canadian government is using artificial intelligence to manage migration. Not much is known about the decision to introduce automated decision-making systems to triage TRV applications in Canada, and therefore, the question that this paper seeks to address is twofold: what, if any, consideration, did the Canadian government give to the risks associated with deploying artificial intelligence in migration management? Furthermore, are the safeguards introduced and steps taken by IRCC in the development of the TRV pilot effective at establishing a positive precedent for testing and subsequently introducing new technologies to deliver migration programs and government services more broadly?

This paper begins with a section dedicated to defining several introductory concepts relating to artificial intelligence and machine learning. What follows is a summary of the research methodology employed and a review of existing literature that acknowledges efforts made to establish clarity on how the Canadian government uses artificial intelligence to manage migration to date.

The paper includes a section that defines the TRV stream and how applications are reviewed and decisions are rendered. Given that the TRV stream is the business line which IRCC has selected to pilot artificial intelligence capabilities, the paper then provides a thorough description of the automated decision-making system introduced into the TRV stream, focusing on the reasons the TRV stream was selected, how the technology was developed and deployed, the quality assurance and accountability systems in place and the role the system has played in the development and use of future artificial intelligence projects. Then, I provide a critical appraisal of the safeguards IRCC introduced in the development and deployment of the TRV pilot, including engaging a multidisciplinary team to develop the system, engaging experts to assess the system's methodology before its launch and selecting a point of intervention that involves consistent, repetitive and frequent decision-making. I review the humans-in-the-loop component of the pilot and the effort to avoid using black-box algorithms. To close, I offer additional considerations for the Canadian Government on the TRV pilot, assessing accountability, methods for monitoring data and trends, the importance of transparency, completing an Algorithmic Impact Assessment and being intentional in the development and use of technologies in migration management.

Relevant Terminology

Artificial intelligence is rapidly becoming a key consideration in many organizations and their digital transformation plans. Despite being the buzzword, artificial intelligence is associated with and involves many other key terms and concepts. Below, I outline relevant terminology used

by experts and in common conversation to establish a framework to better understand this paper. In describing and defining these terms, I also acknowledge some potential shortcomings and opportunities that will be expanded upon throughout the paper regarding migration and asylum governance.

Big Data

The term big data refers to data that is too large, fast or complex to process using traditional methods (SAS, 2021). Big data has several dimensions, including, volume, velocity and variety. Volume refers to the size or amount of information produced from sources – things such as business transactions, Internet devices, videos and social media (SAS, 2021). Velocity is the speed at which data is produced, used and managed, which together create a cycle. Finally, variety refers to the different formats that data arrives in, usually structured and unstructured. Structured data relates to numeric data saved in traditional databases, while unstructured data refers to emails, audio, videos, financial transactions and text documents (SAS, 2021). At this time, for machines to be able to learn (machine learning), the data given to the machine needs to be in the same format: only pictures, only audio files, only text.

The SAS Institute introduces two other dimensions to big data: variability and veracity (SAS, 2021). Variability refers to the varieties of data and data sources, while veracity refers to the quality of data, which is based on the ability to link, match, cleanse and transform information (SAS, 2021). As described below, big data is required to code programs and for machines to learn. Big data can be acquired through crowdsourcing, public and private sources.

Artificial Intelligence

Artificial Intelligence, or 'AI,' refers to "machine-based operations that mimic human intelligence" (Schmidt and Stephens, 2019, p. 133). To better understand AI, it is helpful to view intelligence as a property that something can have: in other words, humans can be intelligent, so too, can machines. However, AI is often viewed as an unfit term to define what advanced technologies are currently capable of, as AI is currently only capable of engaging one component of intelligence – prediction (Agarwal et al., 2018).

Machine Learning, Input Data, Training Data and Feedback Data

To understand prediction in AI, one must first understand machine learning. Machine learning is a subfield of AI that employs mathematics, statistics and computer science to acquire knowledge, and therefore, make predictions effectively (Schmidt and Stephens, 2019). Machine learning automates analytical modelling. A machine, or computer, is assigned a task to complete or a problem to solve. The machine is then given an algorithm (or, mathematical equation), which it uses to complete the task or solve the problem. The machine then repeats the same task or solves the same problem. Over time, this repetition reinforces what the machine has learned to do, and the machine becomes better at completing the task or solving the problem.

The way a machine learns depends on the algorithms it uses. Machine learning can occur in four particular ways, supervised learning, unsupervised learning, semi-supervised learning and reinforcement machine learning (Lorberfeld, 2019). However, for this paper, I will briefly explain supervised and unsupervised learning. Usually, straightforward tasks are done through supervised machine learning, where an individual submits a "training dataset," which includes an input variable and an output variable. The machine then uses an algorithm (mathematical

equation or function) to map a path from the input variable to the output variable. Supervised learning allows the machine to learn from the training dataset so that when there is new input data, the machine can predict what the output variable will be (Brownlee, 2016). This can be demonstrated using the following function, where (x) represents the input variable, (y) the output variable and "f" the mapping function:

$$\mathbf{Y = f(X)}$$

Within supervised machine learning, there are two subcategories: classification and regression (Agarwal et al., 2018, 101). While both classification and regression rely on training datasets, under classification, machines learn to assign the input data as observations in categories, whereas under regression, the output of the above function is a numerical figure (Garbade, 2018).

In contrast, unsupervised machine learning involves finding patterns or mapping functions in data that is not "labelled" – where there is no output variable (Y) available. Under unsupervised learning, the machine uses the algorithm to discover and present interesting structures in the data (Brownlee, 2016). The two subcategories under unsupervised learning are clustering, which involves grouping observations based on similarities, and association, which involves discovering rules that describe large parts of the data. Under clustering, for example, a machine would group customers based on their purchasing behaviour, while under association, the machine defines the rules that describe clustering: customers who purchase eggs also purchase bacon.

Predictive Analytics

As mentioned earlier, to understand AI, it is important to first understand prediction. Both supervised and unsupervised machine learning use prediction. Prediction is "the process of filling in missing information," using information that exists to generate information that does not exist (Agarwal et al., 2018, p. 13). Accordingly, predictive analytics is an area in statistical science that extracts existing data and processes it to forecast – or make predictions about – trends and outcome patterns (Pedamkar, n.d.). Predictive analytics, unlike machine learning, is heavily informed by statistics. The more a machine is used in the machine learning process, the more data it collects and, subsequently, by training on this data, the machine becomes better at making predictions. Today, everything can be seen as a prediction problem: taking information that we have about a problem and using it to discover more. This is exemplified by language translation: once a task for a linguist, is today, a task for Google Translate.

Deep Learning

Deep learning is a branch of machine learning that allows computers to perform human tasks, like understanding a language or identifying images, which uses algorithms inspired by the structure and function of the human brain (Brownlee, 2019). Deep learning employs "artificial neural networks" to undertake both supervised and unsupervised learning methods. In the human brain, neural networks are made up of neurons that process information received from our senses to deliver an output or response (Panchal, 2018). Put simply, this is the process of collecting information from a human sense, interpreting it and storing it in the human "memory." For example, an individual who boils water for the first time and touches the hot pot. The individual's resulting pain is then stored in their memory, they have learned, and upon boiling water again, the individual recalls the associated pain of touching the hot pot.

Deep learning adapts the biology behind the human neural network to allow machines to learn. In deep learning, a machine uses artificial neural networks which are algorithms inspired by the human brain, to progressively develop a better understanding of the relationship between the input (X) and output (Y) data. Just as humans learn over time, in deep learning, algorithms are repeated to perform a task and each time, the method to complete the task is changed slightly to improve the output (Marr, 2018). Deep learning requires big data to learn from. Deep learning is used in processes such as facial and iris recognition. More specifically, in these instances of image recognition,

> deep learning algorithms will recognize edges at a certain level, nose at another level, and face at yet another level. With the ability to understand data from the lower level up the chain, a deep learning algorithm can improve its performance over time and arrive at decisions at any given moment in time. (Wu, 2019)

The Black Box Dilemma

AI is often criticized for having a "black box" problem, where some kinds of machine learning make it difficult if not entirely impossible to understand what machines are doing when they are learning (Kuang, 2017; Bathaee, 2018). According to Yavar Bathaee, the black box problem can be divided into two categories: strong black boxes and weak black boxes (2018). The strong black box is AI that is entirely opaque to humans. In other words, humans are unable to understand how the AI arrived at a decision, prediction or output, what the AI deemed essential in determining the outcome and how the AI ranked the variables it processed in order of importance (Bathaee, 2018). Notably, even after the AI produces an output, the process it used to arrive at the output cannot be analyzed. In contrast, in weak black boxes, the decision-making processes are also opaque to humans. However, unlike strong black boxes, weak black boxes can be probed to understand how the AI ranked the importance of the variables it considered, which may then allow for a limited ability to speculate how the AI arrived at its decision (Bathaee, 2018). Conversely, white boxes are interpretable models, where their behaviour, how they produce predictions and what variables influence decision-making are entirely discernable. Two key elements of a white-box model are features that are understandable and a machine learning process that is transparent (Sciforce, 2020).

Research Methodology

In 2018, Petra Molnar and Lex Gill, at the University of Toronto, released a report indicating that the Canadian Government has been experimenting with the use of algorithms and AI in decision-making related to Canada's immigration and refugee system (Molnar and Gill, 2017). Molnar and Gill had submitted 27 separate Access to Information Requests to the Canadian government, including to IRCC and the Canada Border Services Agency, Canada Security Intelligence Service, Shared Services Canada, Public Safety and Emergency Preparedness, Global Affairs Canada, Innovation, Science and Economic Development, and the Royal Canadian Mounted Police (2017). Molnar and Gill wanted to determine if and what AI technologies were being used by the Canadian Government. While many of their Access to Information Requests remain outstanding, Molnar and Gill discovered that since 2014, IRCC has been developing a system to "identify the merits of an immigration application, spot potential red flags for fraud and weigh all these factors to recommend whether an applicant should be accepted or refused" (Keung, 2017). In their report, Molnar and Gill argue that using automated decision-

making technologies in place of human judgment has the potential to violate domestic and international human rights laws (Kenyon, 2018).

In 2018, IRCC alongside the Department of Justice advertised a Request for Information for a proposed "Artificial Intelligence Solution" to

> understand how each respective department could leverage the benefits of AI, machine learning and predictive analytics technologies to support decisions involving legal considerations… the development of litigation strategies, and generate new insights and efficiencies in the delivery of legal services and litigation support to the Canadian government. (Public Works and Government Services Canada, 2018)

To date, IRCC and the Canadian government have not publicly shared information on how the Canadian government uses AI to manage migration.

To fill the gaps that remain outstanding on how exactly the Canadian government is or intends on using AI to manage migration, this study employed a qualitative research approach by way of semi-structured interviews with individuals employed by IRCC and the Immigration and Refugee Board (IRB). This approach was selected as it offered the opportunity to expand on, strengthen or refute some of the conclusions made regarding the subject matter in the field thus far. To recruit participants, I invited individuals who indicated on their LinkedIn[1] profile that they were employed by IRCC or the IRB to connect. With the invite, I sent a personalized 300-character note providing a summary of my research study. I employed a snowball effect to recruit further participants through individuals I had already contacted.

In total, I contacted 198 individuals via LinkedIn. Of the 198 individuals I contacted, 89% did not agree to participate in an interview, of which 33% indicated that they were uncomfortable participating in a study out of fear of sharing confidential information and 56% read about the research study and declined the invitation to connect. While not central to the theme of this paper, the number of individuals who indicated an unwillingness to participate in an academic study related to public service is rather high. Although many factors may have contributed to individuals' decision to decline to participate, of the 89% that declined, 33% of individuals specifically declined out of fear of sharing confidential information. This speaks to the disconnect between academia and public service that many public administration scholars often cite as an issue in assessing and ultimately improving public service delivery (Raadschelders and Lee, 2011). Despite the rather low number of participants, it should be noted that given that the information I was aiming to collect was administrative, I was able to reach a point of saturation, where different participants were sharing the same information as one another. Individuals who agreed to participate were interviewed via telephone, Skype or Zoom.

Notably, while the central purpose of the interviews was to determine if and how AI was being used in decision-making relating to Canadian immigration applications, public servants also identified many strengths, shortcomings and opportunities with the technology that currently exists to manage migration. These insights are captured throughout this paper.

The Temporary Resident Visa

Often referred to as a visitor visa, the temporary resident visa (TRV) is one avenue for immigrating to Canada. A TRV is "an official document issued by a Canadian visa office in [an

[1] I selected LinkedIn for two reasons: the contact information of Canadian public servants in the departments of interest are not publicly accessible and LinkedIn offered a more informal method of communication.

individual's passport] to show that [they] have met the requirements for admission to Canada as a temporary resident (either as a visitor, a student, or a worker)" (Government of Canada, 2020a). As per section 22 of the *Immigration and Refugee Protection Act (IRPA),*

> ***Temporary resident***
> *22 (1) A foreign national becomes a temporary resident if an officer is satisfied that the foreign national has applied for that status, has met the obligations set out in paragraph 20(1)(b), is not inadmissible and is not the subject of a declaration made under subsection 22.1(1).*
>
> ***Obligation on entry***
> ***20 (1)*** *Every foreign national, other than a foreign national referred to in section 19, who seeks to enter or remain in Canada must establish,*
> […]
> *(b) to become a temporary resident, that they hold the visa or other document required under the regulations and will leave Canada by the end of the period authorized for their stay.*

Individuals aged 18 or older applying for temporary residency in Canada are required to complete the Application for Temporary Resident Visa [IMM 5257] and Family Information [IMM 5707] IRCC forms. Additional forms that may need to be completed include the Statutory Declaration of Common-Law Union [IMM 5409], Use of a Representative [IMM 5476] and Authority to Release Personal Information to a Designated Individual [IMM 5475] forms (Government of Canada, 2020b). Alongside these forms, individuals are required to pay the processing and biometric collection fee and submit the required supporting documents (such as a valid passport) to accompany their application.

Eligibility, Admissibility and Refusal

TRV applications are reviewed for both eligibility and admissibility. Eligibility refers to whether an applicant meets the requirements of the immigration program they have applied for, things like age, work experience and education, language proficiency, adaptability and arranged employment. Admissibility refers to the applicant's security, criminal, medical and financial information. These factors remain the same for all streams of immigration. According to Canadian immigration law, specifically under Division 4 under Part 1 of the *Immigration and Refugee Protection Act,* the factors related to admissibility include security, human or international rights violations, serious criminality, criminality or organized criminality, health grounds, financial reasons or misrepresentation, amongst other reasons (IRPA S.C. 2001, c. 27). Notably, this particular section of the IRPA was drafted with discretion being written straight into it, meaning the list is open-ended, allowing for considerations not previously enumerated.

In instances where applicants are found ineligible, the decision to refuse the applicant may be made after review of a written application with no in-person interview conducted with the applicant or, after an interview with the applicant (Government of Canada, 2018). If a review of the paper application makes evident that the applicant is ineligible and additional information would not alter a refusal decision, the applicant will not be invited for an interview. An application may be refused if the applicant does not meet the obligations of the *Immigration and Refugee Protection Act* and its Regulations, including subsection A11, paragraph A20(1)(b) and subsection R179. Additionally, an application may be refused if the visa officer is not satisfied the applicant will leave by the end of the period authorized for their stay or any grounds for inadmissibility as per the *Immigration and Refugee Protection Act,* including security, human or international rights

violations, criminality, organized crime, health, financial reasons, misrepresentation, non-compliance with the Act, or, inadmissible family member (Government of Canada, 2018). Refused applicants may seek redress from the Federal Court of Canada and the Canadian Human Rights Commission.

Introducing Automated Decision-Making to Triage TRV Applications

Selecting the TRV Stream

As outlined in Figure 1, Canada has seen a steady increase in TRV applications from 2012 to 2020.[2] This increase and subsequent greater demand for resources to process applications lead IRCC to consider novel methods in increasing the processing of applications, ultimately resulting in the introduction of a pilot project to implement advanced analytics and automation to decision-making relating to client applications. IRCC was driven to develop and use the aforementioned system in an attempt to increase efficiency, develop a sustainable method to respond to and process the increasing volume of applications and provide better client service. IRCC involved a multidisciplinary team with experts of different backgrounds to explore and develop the technology to ensure the proper support for the delivery of services. The pilot project, titled "TRV eApps Advanced Analytics Pilot" (henceforth the "TRV pilot"), automates a portion of the TRV business process, strictly focusing on online applications from China and India.

IRCC decided to test AI capabilities on the TRV stream for four particular reasons: complexity, stakes, need and data. More specifically, the TRV stream is not particularly complex: while certain applications involve routine examination of information, other cases require further, more nuanced judgment. The stakes are relatively lower for TRV applicants if they are denied entry into Canada[3]; in other words, individuals applying for a visitor visa face lower risks than individuals applying for permanent residence, where the outcome of the application has the potential to change the course of a person's life. Finally, the Chinese and Indian online TRV applications offer greater amounts of structured data. IRCC decided to pilot the automated decision-making system to triage applications from China and India because these countries were where the majority of applications in the TRV stream were being made.

Developing and Deploying the Technology

Having established the TRV stream as the focus of the pilot project, IRCC's data governance team developed hundreds of models to triage applications before selecting the best one. Each model was trained and metrics were used to measure each algorithm's performance. Sufficient data for training the models were selected through an intensive process to represent the TRV business line. In particular, data was selected from TRV applications that were approved from 2015-2017 – and the data was then reviewed for quality and completeness. Low quality or incomplete applications were excluded from the dataset that the models were trained on. To ensure the model rules were representative of recent trends, the models developed for China and

[2] Data from 2017 is currently unavailable. The author is in the process of requesting and obtaining this data and this section will be updated when the corresponding information is received.

[3] Notably, how the stakes of an application are perceived depends on an individual's reasons for applying for a TRV. Some applicants would not describe their TRV application as low-stakes. Take, for example, an individual who would like to visit their dying family member or meet an important business client. While a TRV may be required, the outcome of being denied a TRV may weigh heavily on or 'cost' a lot for an individual who cannot visit their dying family member or make it to the meeting.

India were built based on applications recently received from those countries. Before launching the first system, the National Research Council of Canada conducted an expert assessment of the methodology developed by IRCC and concluded that

> the methodology of the initiative is excellent and follows the steps necessary for the success of a machine learning project. The initiative is very well adapted to the various organizational risks (legal, public perception, security) while maximizing performance measures, i.e. transparency of methodology, efficiency of operations and data quality. (Government of Canada, 2020c, p. 5)

Advanced analytics and automation have been implemented in the eligibility phase of the TRV application review process. The model applies rules established by experienced visa officers ("officer rules") to determine whether applications are complex and should be reviewed and decided upon according to regular procedures by a visa officer, or whether they can be fed through the model. The officer rules identify characteristics in applications indicating a higher likelihood of ineligibility or inadmissibility (for example, criminality or security threats) or which require additional documents (like a consent letter for a child travelling alone) (Government of Canada, 2020c).

As per Figure 1 applications that do not trigger the complexity indicator based on the officer rules in the first review are then fed into the system based on rules developed by a machine learning algorithm. The system triages applications into three tiers. Tier 1 applications are the simplest (for example, a returning applicant who has a positive travel history) and therefore, eligibility is approved by the system solely based on the model's determination, without review by the visa officer. Tier 2 and 3 applications are of medium and high complexity, respectively, and therefore, are assigned to visa officers to make determinations on both eligibility and admissibility. The model does not make any decisions relating to either eligibility or admissibility for Tier 2 or 3 applications, which are decided upon by visa officers. Visa officers then review all applications for admissibility and render a final decision. To be clear, a human reviews all TRV applications. The decision-making process is only partly automated and the final decision is never rendered by the automated decision-making system.

Quality Assurance and Accountability

Each day, 10% of Tier 1 applications assessed for eligibility by the model are given to visa officers to determine the eligibility and admissibility of the applicant. This is a 'blind' quality assurance method: the visa officers are unaware that the applications they are reviewing have been deemed eligible by the model. Once visa officers have decided on the eligibility of the applicants, their decisions are compared with those made by the model. IRCC conducts this blind quality assurance test daily to ensure the Department retains a 99% concurrence rate between the visa officers and the model's approval of eligibility. From the date the pilot was deployed, until February 19, 2020, the model has met a 99% concurrence rate. Notably, the process is a 'humans-in-the-loop' model: the model does not simply make a decision that is accepted as true and correct. Rather, experienced decision-makers vet the rules that the algorithm is trained on to make them fit the day-to-day best practices and heuristics visa officers would otherwise use in the manual review process.

Other quality assurances and monitoring measures include tracking the daily number of applications triaged to each tier to ensure the model is functioning as intended and outages or disruptions do not occur in IRCC's IT system. Additionally, IRCC monitors trends to determine if the model rules should be updated or the model must be retrained. Notably, IRCC tracks the adverse behaviour of applicants admitted to Canada to determine if there must be changes or

reconsiderations in how, where and why the model is used. To ensure accountability in the use of these systems, visa officers reviewed and approved each of the business rules developed for the model to triage applications into Tier 1. A Director General then provided a formal approval of the officer rules previous to their implementation in the system. Upon approval, this Director General became the decision-maker of record for all eligibility approvals issued through the system. This is done to ensure that lines of accountability are clearly defined and lead to a human decisionmaker.

A Role in Future AI Projects

The TRV pilot was launched before the development of Canada's Directive on Automated Decision-Making. The Directive, which came into force on April 1, 2020, is used to ensure the development and use of AI complies with administrative law principles. For systems in place before this date (like the TRV pilot), the Directive serves as a best practice guide. Systems developed after this date must comply with all of the Directive's requirements. The Directive has several requirements for the individual responsible for the program using the automated decision system, including quality assurance, the completion of an algorithmic impact assessment, transparency, recourse, and reporting.

In a special lecture delivered to the Law Society of Ontario in 2019, IRCC explained that because its TRV pilot was one of the most concrete examples of AI in government at the time, the department directly engaged with and provided feedback to the Treasury Board Secretariat of Canada in the development of the Algorithmic Impact Assessment tool accompanying the Directive. The Directive on Automated Decision-Making stipulates that before producing an automated decision system, the Deputy Head must complete an Algorithmic Impact Assessment, assessing and mitigating the risks associated with deploying the automated decision system. Although the Directive was developed and published after the launch of the TRV pilot, IRCC has indicated it "is prepared to demonstrate full compliance with the Directive once it comes into force."

The Safeguards Introduced in IRCC's TRV pilot

Through the introduction of the TRV pilot in 2018, IRCC became one of the earliest adopters of AI in the Canadian government. However, this recent development raises significant concerns about the ramifications of using AI in Canada's immigration system. IRCC has introduced some safeguards in the development and deployment of the TRV pilot, including engaging a multidisciplinary team to develop the system, engaging experts to assess the system's methodology before its launch and selecting a point of intervention that involves consistent, repetitive and frequent decision-making. Additionally, the humans-in-the-loop component of the pilot and the effort to avoid using black box algorithms are key features of the pilot that have the potential to safeguard against errors. Further consideration should however be given to these safeguards and their effectiveness, as well as how accountability is delineated and how data and trends are monitored.

Engaging a Multidisciplinary Team to Develop the System

According to IRCC,

> a diverse, multi-disciplinary team of experts has been engaged in problem-solving – no longer compartmentalizing issues and tackling each separately, but instead

> working on the whole thing at once: the technology and data science, the business process, the legal issues, the policy questions. (McEvenue and Mann 2020, p. 3)

Engaging a team of individuals who have particular expertise in a diverse range of relevant subjects is incredibly important in developing an automated decision-making system as it offers diverse perspectives when considering complex problems related to the development and deployment of AI systems. Additionally, a multi-disciplinary team offers the ability to detect biases that may infiltrate the algorithms.

A good next step for IRCC and the Canadian government would be to publicly clarify who has been and is involved in the development and use of the TRV pilot. In government, and public administration, there is a general understanding that consulting stakeholders is an important step in the policy, program and service development and delivery process. This is done to ensure that stakeholders can see themselves reflected in the policies that impact them and that programs and services are designed with the user in mind. However, concerning the migration sector, sharing information is as important if not increasingly more important than consultations. This is because immigrants – the users of these policies, programs and services – are not afforded the same rights to weigh in on government decisions and provide feedback and therefore, come to rely on community groups to advocate on their behalf. Therefore, while having a multi-disciplinary team develop the technology is a good first step, to continue on the right path, the Canadian government should be transparent about the TRV pilot and who has been involved in its development. Sharing information about the TRV pilot with external partners that advocate on behalf of users of the TRV stream will offer the Canadian government an opportunity to question its assumptions and get feedback on what the overall user journey is like and where the TRV system can be improved.

Engaging Experts to Assess the System's Methodology Before Launching

In 2018, before launching the initial system, IRCC requested the National Research Council of Canada to conduct an expert assessment of its methodology. The National Research Council of Canada agreed and concluded that

> the methodology of the initiative is excellent and follows the steps necessary for the success of a machine learning project. The initiative is very well adapted to the various organizational risks (legal, public perception, security) while maximizing performance measures, i.e. transparency of methodology, efficiency of operations and data quality. (Government of Canada, 2020c, p. 5)

While on the surface this assessment sounds compelling, it is important to note that the National Research Council of Canada is itself a Canadian government organization. Therefore, despite any effort to remain free of bias, the National Research Council of Canada's assessment was effectively the Canadian government assessing itself.

To encourage transparency and accountability and gain a third-party review of the TRV pilot, a good next step for the Canadian government would be to request an external body with no vested interest in the success of the TRV pilot to assess the pilot project and identify any strengths and deficiencies.

Launching the System in a Consistent, Repetitive and Frequent Decision-Making Process

Introducing the TRV pilot into the eligibility portion of the TRV triage for online applications received from China and India was a thoughtful decision. Captured in Figure 3, as IRCC itself has indicated, China and India have consistently made up a large number of the applications received through the TRV stream, ranging from at least 33.47% in 2012, to at most, over half of the applications, at 51.74% in 2018. What is most notable, however, is that from 2012 to 2018, the top-ranking countries from which TRV applications have come from have been India and China, respectively. Therefore, IRCC's reasoning for introducing the TRV pilot to triage applications for eligibility from India and China is rational, as the Chinese and Indian online TRV applications would offer greater amounts of structured data relatively compared to other applications, particularly as these countries were where the majority of applications in the TRV stream were coming from and being decided upon.

The Nuanced Details of Keeping Humans-in-the-Loop

"Humans-in-the-loop" refers to the involvement of real humans in the development and operation of machine-learning systems. By leaving all final decisions to be made by a human, IRCC takes an appropriate step to ensure meaningful human agency in decision-making processes. However, one notable concern that must be assessed concerning the role of humans – in this case, visa officers – as they interact with and use or rely on the TRV pilot is that of control and complacency. As per John Zerilli, Alistair Knott, James Maclaurin and Colin Gavaghan, there is a control problem that arises when the human in "a human-machine control loop… become[s] complacent, over-reliant or unduly diffident when faced with the outputs of a reliable autonomous system" (Gavaghan et al., 2019, 555).

In the context of the TRV pilot, visa officers are instructed not to revisit the eligibility decision on Tier 1 applications unless information affecting the positive eligibility of the applicant is brought to the visa officers' attention in conducting the admissibility assessment or when verifying the applicant's passport. In other words, beyond the point at which visa officers review a random sample of Tier 1 applications each day to determine the eligibility and admissibility of the applicant, unaware that the eligibility portion of some of those applications had been decided upon by the TRV pilot, decisions rendered by the TRV pilot are not to be reviewed by visa officers. This course of action allows for error to occur as a result of proactive complacency. In other words, by not allowing visa officers to review decisions rendered by the TRV pilot, IRCC may encourage officers to accept the TRV pilot's decision as correct without certainty.

It is unclear at this time whether IRCC shares progress or update reports on the TRV pilot and its performance internally with visa officers. If officers are aware that the TRV pilot boasts a relatively high concurrence rate (which is an otherwise positive metric), this may reinforce the behaviour to accept the TRV pilot's decisions as accurate without review.

Avoiding the Black Box

A member involved with the development of the TRV pilot indicated that, in developing the system, IRCC officials found that the performance gain of the black box algorithm was not particularly large compared to the white box algorithm, and therefore, IRCC selected a white box algorithm, where the machine learning model is transparent and the features of the system are understandable. The decision to select a system where the decision-making process is not only transparent but also understandable was one of the greatest safeguards that IRCC introduced in the TRV pilot. Being able to understand why a decision was made is a particularly important feature of Canada's immigration system, especially given that individuals whose TRV applications are refused are allowed to seek recourse, where case notes outlining the circumstances of the

application, the process followed in coming to or making the decision, the presence and identity of an interpreter, if applicable, representations made by interested persons (or counsel) and the nature and content of these representations as well as the detailed reasons for the refusal are assessed (Government of Canada, 2018). By opting for a white box algorithm, IRCC ensured that how the TRV pilot arrives at a decision to refuse an application based on eligibility can be understood, allowing for a thorough assessment of the decision.

Additional Considerations and Concluding Remarks

While the Canadian government appears to have introduced multiple safeguards to ensure that the TRV pilot is not only an effective but responsible, AI initiative, there remain many outstanding questions that make this assessment difficult. While more information is certainly required to conduct a complete assessment of the pilot, some additional considerations exist that the Canadian Government would benefit from when iterating the TRV pilot. These include a reassessment of accountability in the use of the TRV pilot, clarity on how data and trends are monitored, transparency to encourage advocacy, completing an Algorithmic Impact Assessment and being intentional in the development and use of technologies in migration management.

Loose Lines of Accountability

In advance of introducing the TRV pilot, visa officers validated the officer rules that lead to applications being triaged into Tier 1. These rules have been approved by a Director General, who upon approval, became the decision-maker for all eligibility approvals issued by the system. This is done to ensure that accountability is clearly defined and rests with a human being. For individuals who submit an application and are subsequently unsatisfied with the decision they receive, the same recourse mechanisms apply as before the introduction of these systems.

Currently, there is insufficient information publicly available that clarifies if and how visa officers or the Director General are involved in the process when an individual who is unsatisfied with the decision they receive seeks recourse. It remains to be seen who will be involved in these procedures and how IRCC will address the involvement of an automated decision-making system in the TRV stream. Regardless, a worthwhile consideration is the distance between the decision rendered by the TRV pilot and the Director General which becomes the decision-maker for all of the system's eligibility approvals. As previously mentioned, eligibility in this context relates to whether an applicant meets the requirements of the immigration program they have applied for, things like age, work experience and education, language proficiency, adaptability and arranged employment. While these considerations are relatively less consequential than those that fall within the admissibility assessment that an officer decides on, it is unclear why the Director General becomes the individual accountable for the TRV pilot and how it triages applications.

Monitoring Data and Trends to Ensure the Algorithm is Up-to-Date

Another component of the TRV pilot that, on the surface, appears as a promising effort to ensure the TRV pilot functions as intended is IRCC's indication that the department conducts periodic monitoring over longer periods. Notably, the data that the model was developed on was from 2015-2016, roughly five to six years ago. It is unclear if and how often IRCC iterates the data or what trends are assessed to determine whether the algorithm needs to be retrained and how often re-assessment and re-training occur. It is difficult, however, to determine what “monitoring” in this context entails as IRCC has not, at this time, shared information to outline what trends are

assessed to determine whether the algorithm needs to be retrained and how often the re-assessment and re-training occurs.

Additionally, it is also unclear how IRCC tracks the behaviour of clients, if and where this data is stored and how it may be used to re-train the algorithm. A lack of clarity around these monitoring endeavours makes it difficult to conduct an informed assessment and raises questions about how improvements are made to the pilot and what other information is tracked by the state that loops back into the TRV pilot.

Being Transparent in the Face of the Paradox of Immigration Policy

Information on how exactly AI is used in Canada's immigration system is difficult to locate. My conversations with IRCC officials were fruitful; however, without them, it would be difficult to ascertain if AI was exacerbating disparities within Canada's immigration system as they exist. While officials have expressed an interest in continuing to engage with members of the public, professionals, experts and academics at the intersection of AI and immigration to encourage transparency, a more effective approach would be to make information about the TRV pilot publicly available. By being transparent about the use of advanced analytics and AI to manage migration, IRCC and the Canadian government would greatly benefit from the perspectives of experts that have not yet been engaged.

The role of transparency in the context of piloting AI solutions in immigration is particularly important due to the paradox of immigration policy. An important practice in the development of programs, policies and services is user research – engaging the individuals, groups and communities who are impacted by and may benefit from the relevant policies, programs and services. However, because states develop their own immigration policies, foreigners seeking to immigrate to a country are unable to participate in and help form the policies and programs that might allow them entry into said country. Accordingly, as mentioned above, by making information about how AI is used in the TRV pilot publicly accessible, the Canadian government offers external partners, including non-profit organizations and advocacy groups, the opportunity to provide feedback and recommendations on the strengths of the system and how it can be improved.

Completing the Algorithmic Impact Assessment

As per the Directive on Automated Decision-Making, automated decision-making systems that were introduced before April 1, 2020, the Directive serves as a best practice guide and only systems developed after this date must comply with all of the Directive's requirements. Although the Directive and the accompanying Algorithmic Impact Assessment were developed and published after the launch of the TRV pilot, having played a significant role in the development of the Algorithmic Impact Assessment, IRCC should make good on its own statement and reinforce the government's commitment to administrative law and the responsible development and use of AI by demonstrating full compliance with the Directive.

Being intentional in the development and use of technologies in migration management

It remains unclear from the information that is publicly available whether IRCC conducted user interviews with visa officers to understand whether an automated decision-making system would be the best approach in assisting visa officers in the work they are expected to do. The Department has been clear that it established an internal steering committee to serve as a focal

point for policy, legal, and privacy issues related to IRCC's use of technology and that the decision to introduce AI into the TRV stream was one motivated by an effort to increase efficiency, develop a sustainable method to respond to and process the increasing volume of applications and provide better client service. As the Canadian government endeavours to introduce new technologies and modernize old systems, like the Global Case Management System (GCMS), IRCC's "single, integrated and worldwide system used internally to process applications for citizenship and immigration services" (Government of Canada, 2012), it must be cognizant of the need for interoperability and an integrated system that allows all of IRCC's assets to work together.

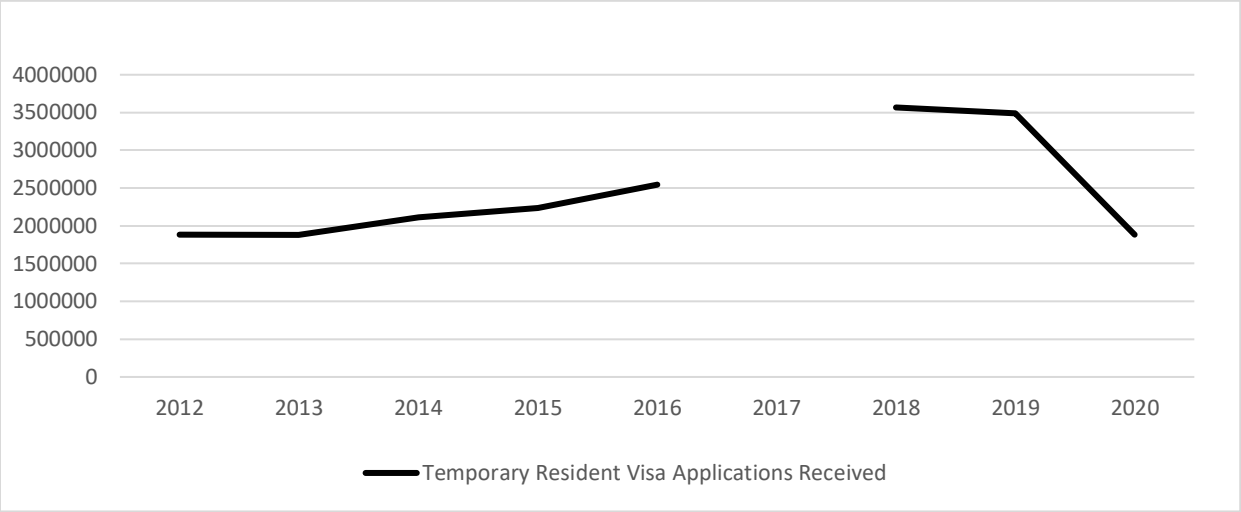
The lack of interoperability of systems is something that many public servants at IRCC and the Immigration and Refugee Board have pointed to concerning the GCMS and legacy systems. Conversations with public servants at IRCC and the Immigration and Refugee Board revealed that the Global Case Management System (GCMS) is accessible by the Canada Border Services Agency, embassies and agencies in other countries where individuals can apply to come to Canada. The GCMS is used to store all information about individuals seeking to come to Canada, including biometric data, notes from interviews, passport information, potential warnings or flags submitted by the Canada Border Services Agency among other information. One public servant indicated that IRCC stores data online, as scanned documents or PDFs, as entries in the GCMS or Excel spreadsheets and physical files in its offices. In modernizing current systems and introducing new systems, IRCC should consider engaging in user research and engage systems thinking to better understand immediate needs, strike a balance between short-term and long-term initiatives and their associated opportunity costs and ensure that systems will be built with interoperability in mind. Before introducing any further systems, IRCC should consider developing a clear digital strategy that engages strategic foresight, "a structured and systematic way of using ideas about the future to anticipate and better prepare for change" (Organisation for Economic Co-Operation and Development, n.d.).

An assessment of some of the safeguards introduced in IRCC's TRV pilot shows that the Department has taken some steps in the right direction to ensure the proper use of AI in delivering the TRV program. However, a lack of sufficient publicly accessible data, both qualitative and quantitative, limits the scope of this analysis. A central aim of this paper has been to first establish and bring to light how the Canadian government is using artificial intelligence to manage migration. A review of the automated decision-making system that IRCC has introduced to triage TRV applications received online from both India and China reveals that while IRCC has given thoughtful consideration to the potential challenges posed by artificial intelligence, more can be done to ensure the pilot is effective. The Canadian government strategically selected the TRV stream as the business line to pilot artificial intelligence capabilities and was careful in how it developed and deployed the technology.

The quality assurance and accountability systems in place are promising: engaging a multidisciplinary team to develop the system, engaging experts to assess the system's methodology before its launch and selecting a point of intervention that involves consistent, repetitive and frequent decision-making are thoughtful and appropriate efforts, and the humans-in-the-loop component of the pilot paired with the decision to avoid using black box algorithms shows that the Canadian government and IRCC are well-intentioned. With some additional investments to clarify accountability, a re-assessment and transparency on the methods for monitoring data and trends to iterate the model, bringing transparency to the system and completing an Algorithmic Impact Assessment, the Canadian government can emerge as a leader in the use of artificial intelligence, particularly to manage migration. Emerging from the COVID-19 pandemic, countries are preparing to re-open borders and invite immigrants in again. Canada has the perfect opportunity to be intentional in the development and use of technologies in migration management to leverage the benefits of emerging technologies while minimizing the risks they can pose to humans.

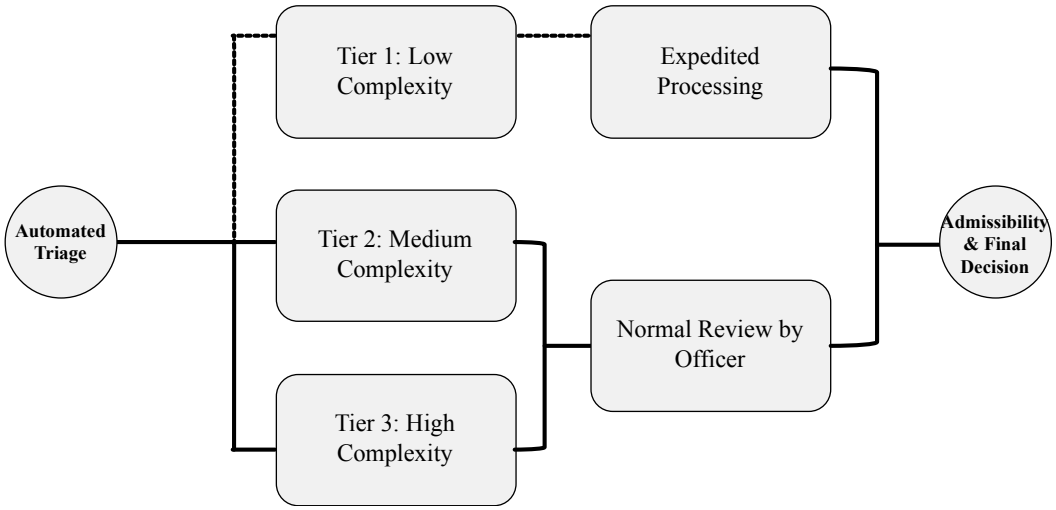
Appendix

Figure 1. Temporary Resident Visa Applications Received Annually from 2012-2020



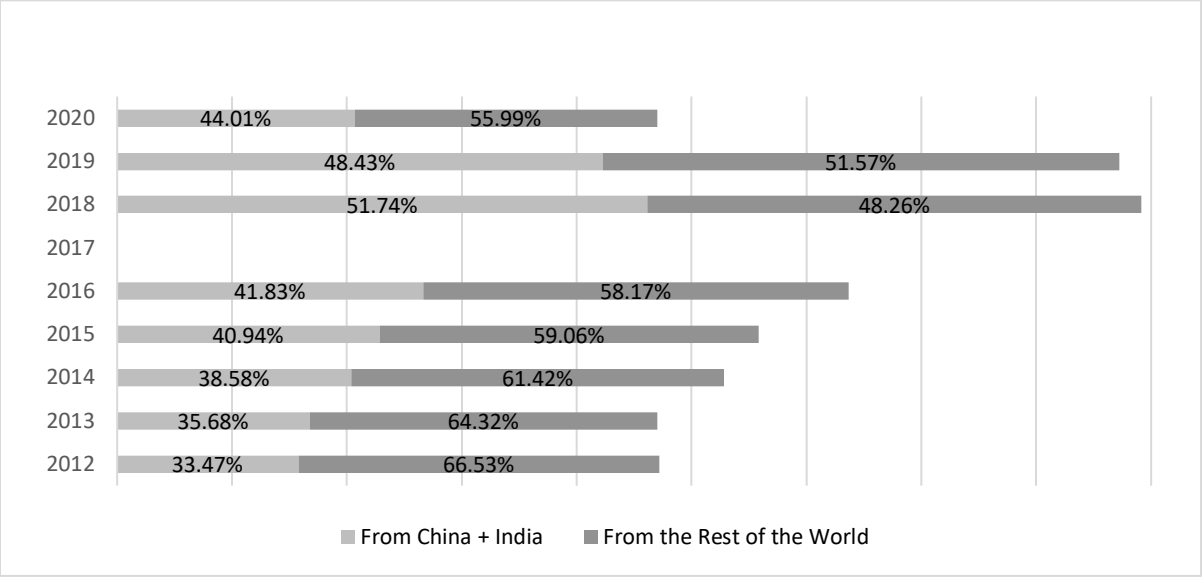
Note - Produced by author. Data for 2012-2016 derived from Government of Canada, 2016b; data for 2018-2020 derived from Government of Canada, 2016a.

Figure 2. TRV Application Processing Using Machine Learning



Note - Produced by author, reviewed and approved by IRCC Strategic Policy and Planning Branch

Figure 3. TRV Applications from China and India vs. the Rest of the World, 2012-2020[4]



Produced by author. Data for 2012-2016 derived from Government of Canada, 2016b; data for 2018-2020 derived from Government of Canada, 2016a.

[4] Data from 2017 is currently unavailable. The author is in the process of requesting and obtaining this data and this section will be updated when the corresponding information is received.

References

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence.* Boston: Harvard Business Review Press.

Bathaee, Y. (2018). The AI Black Box and the Failure of Intent and Causation. *Harvard Journal of Law & Technology 31(2),* 889-938.

Brownlee, J. (2016). *Supervised and Unsupervised Machine Learning Algorithms.* Machine Learning Mastery. https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/

Brownlee, J. (2019). *What is Deep Learning?* Machine Learning Mastery. https://machinelearningmastery.com/what-is-deep-learning/

Garbade, M. J. (2018). Regression Versus Classification Machine Learning: What's the Difference? *Medium.* https://medium.com/quick-code/regression-versus-classification-machine-learning-whats-the-difference-345c56dd15f7

Gavaghan, C., Knott, A., Maclaurin, J., & Zerilli, J. (2019). Algorithmic Decision-Making and the Control Problem. *Minds & Machines*, *29*, 555–578. DOI: 10.1007/s11023-019-09513-7.

Government of Canada. *Immigration and Refugee Protection Act* (S.C. 2001, c. 27).

Government of Canada. (2012). *Privacy Impact Assessment Summary – Global Case Management System (GCMS) - Phase II.* Government of Canada. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/transparency/access-information-privacy/privacy-impact-assessment/global-case-management-system-2.html.

Government of Canada. (2016a). *[ARCHIVED] Source Countries - Applications Received for Temporary Residents,* electronic dataset, Open Datasets and Open Information records. Record ID: *63858380-265e-4948-8e96-ca50fceca2d7*

Government of Canada. (2016b). *Source Countries - Applications Received for Permanent Residency (in Persons) by Month January 2018 to March 2021*, electronic dataset, Open Datasets and Open Information records. Record ID: 9b34e712-513f-44e9-babf-9df4f7256550

Government of Canada. (2018). *Temporary residents: Refusals overview.* Government of Canada. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/publications-manuals/operational-bulletins-manuals/temporary-residents/visitors/refusals-overview.html

Government of Canada. (2020a). *Applying for a visitor visa (temporary resident visa) – IMM 5256.* Government of Canada. https://www.canada.ca/en/immigration-refugees-citizenship/services/application/application-forms-guides/guide-5256-applying-visitor-visa-temporary-resident-visa.html

Government of Canada. (2020b). *Temporary residents: Document requirements.* Government of Canada. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/publications-manuals/operational-bulletins-manuals/temporary-residents/visitors/document-requirements.html.

Government of Canada. (2020c). *Overview of the Analytics-Based Triage of Temporary Resident Visa Applications.* 1-8.

Kenyon, M. (2018). *Bots at the Gate: A Human Rights Analysis of Automated Decision Making in Canada's Immigration and Refugee System.* https://citizenlab.ca/2018/09/bots-at-the-gate-human-rights-analysis-automated-decision-making-in-canadas-immigration-refugee-system/

Keung, N. (2017). *Canadian immigration applications could soon be assessed by computers.* Toronto Star. https://www.thestar.com/news/immigration/2017/01/05/immigration-applications-could-soon-be-assessed-by-computers.html

Kuang, C. (2017). *Can A.I. Be Taught to Explain Itself?* New York Times. https://www.nytimes.com/2017/11/21/magazine/can-ai-be-taught-to-explain-itself.html.

Lorberfeld, A. (2019). *Machine Learning Algorithms in Layman's Terms, Part 1.* Towards Data Science. https://towardsdatascience.com/machine-learning-algorithms-in-laymans-terms-part-1-d0368d769a7b

Marr, B. (2018). *What Is Deep Learning AI? A Simple Guide With 8 Practical Examples*. Forbes. https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/#199695aa8d4b

McEvenue, P., & Mann, M. (2019). "Case Study: Developing guidance for the responsible use of artificial intelligence in decision-making at Immigration, Refugees and Citizenship Canada." *Law Society of Ontario, Special Lectures 2019.* 1-32.

Molnar, P., & Gill, L. (2018). *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System*. Toronto: University of Toronto. https://ihrp.law.utoronto.ca/sites/default/files/media/IHRP-Automated-Systems-Report-Web.pdf

Organisation for Economic Co-operation and Development. (n.d.). *Strategic Foresight.* OECD. https://www.oecd.org/strategic-foresight/

Panchal, S. (2018). *Artificial Neural Networks – Mapping the Human Brain.* Medium. https://medium.com/predict/artificial-neural-networks-mapping-the-human-brain-2e0bd4a93160

Pedamkar, P. (n.d.). *Predictive Analytics vs Data Science.* Educba. https://www.educba.com/predictive-analytics-vs-data-science/.

Public Works and Government Services Canada. (2018). AI Solution (B8607-180311/A), Tender Notice. https://buyandsell.gc.ca/procurement-data/tender-notice/PW-EE-017-33462.

Raadschelders, J. C. N., & Lee, K-H. (2011). Trends in the study of public administration: empirical and qualitative observations from Public Administration Review, 2000–2009. *Public Administration Review, 71 (1),* 19-33. DOI: 10.1111/j.1540-6210.2010.02303.x

SAS. (2021). *Big Data: What it is and why it matters.* SAS. https://www.sas.com/en_ca/insights/big-data/what-is-big-data.html#history

Schmidt, N., & Stephens, B. (2019). An Introduction to Artificial Intelligence and Solutions to the Problems of Algorithmic Discrimination. *Cornell University: Computers and Society,* 130-144. https://tinyurl.com/yhnwfajo

Sciforce. (2020). *Introduction to the White-Box AI: the Concept of Interpretability.* Medium. https://medium.com/sciforce/introduction-to-the-white-box-ai-the-concept-of-interpretability-5a31e1058611#:~:text=White%2Dbox%20models%2C%20or%20in,what%20the%20influencing%20variables%20are.

Wu, J. (2019). *AI, Machine Learning, Deep Learning Explained Simply.* Towards Data Science. https://towardsdatascience.com/ai-machine-learning-deep-learning-explained-simply-7b553da5b960