APPENDIX A

TORONTO POLICE SERVICES BOARD

USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGY

DATE APPROVED		
DATE(S) AMENDED		
DATE REVIEWED		
REPORTING REQUIREMENT	Chief to report to Board from time to time as required by this Policy and directed by the Board.	
LEGISLATION	*Police Services Act,* R.S.O. 1990, c. P.15, as amended, s. 31(1)(c). *Municipal Freedom of Information and Protection of Privacy Act*, R.S.O. 1990, c. M.56. Human Rights Code, R.S.O. 1990, c. H.19. The Canadian Charter of Rights and Freedoms	

Guiding Principles

The Toronto Police Services Board (the Board) supports the efforts of the Toronto Police Service (the Service) and its Members to provide effective and accountable policing through the prudent adoption of new technologies, while, at the same time, ensuring transparency and making certain that policing is provided in accordance with both the law and the interests of the public, and protects and promotes fundamental rights.

Novel technologies making use of artificial intelligence (AI) applications hold the promise of improving the effectiveness of policing services and increasing public safety in Toronto. At the same time, technological advancements may pose new concerns for the privacy, rights (including the rights to freedom of expression, freedoms of association and freedom of assembly), dignity and equality of the individuals affected by them. For example, there have been instances in which novel technologies were shown to incorporate and perpetuate pre-existing and systemic biases, resulting in both individually and systemically discriminating decisions. Furthermore, such unintended consequences may undermine the desired benefits to efficiency and effectiveness of policing services, as well as public trust in policing.

Under section 41(1)(a) of the *Police Services Act* (the *Act*), the Chief of Police is responsible for administering the police service and overseeing its operation, in accordance with the objectives, priorities and policies established by the Board.

The Board is the entity that is responsible for the provision of adequate and effective policing under the *Act* and its successor legislation. No current statutes or regulations fully govern the use of AI technologies in Ontario or Canada, and the Province has not yet developed comprehensive guidelines for the use of such technologies in policing. As a result of the current legal gaps and desired use of AI technologies, the Board determines it necessary to establish governance to facilitate decision-making that is in the public interest, and to enable the Chief to assess and account to the Board concerning how technology will be procured, implemented and used in the provision of policing in Toronto. In its review of proposed AI technologies, the Board will consider the need for and benefits of deploying the new technology; the potential unintended consequences to the privacy, rights, freedoms and dignity of members of the public and Service Members, and to the equitable delivery of police services to the public; and, any possible mitigating actions to eliminate any such unintended consequences. To the greatest degree possible, the Board must conduct such reviews in public.

All use of technology, including AI technology, whether approved by the Board or otherwise, must adhere to the following guiding principles:

- **Legality**: All technology used, and all use of technology, must comply with applicable law, including the *Police Services Act* (and its regulations, as well as successor legislation), Ontario's *Human Rights Code*, and the *Canadian Charter of Rights and Freedoms*, and be compatible with applicable due process and accountability obligations.

- **Fairness**: Use of AI technology must not result in the increase or perpetuation of bias in policing and should diminish such biases that exist.

- **Reliability**: AI technology must result in consistent outputs or recommendations and behave in a repeatable manner.

- **Justifiability**: The use of AI technology must be shown to further the purpose of law enforcement in a manner that outweighs identified risks.

- **Personal Accountability**: Service Members are accountable, through existing professional standards processes, for all the decisions they make, including those made with the assistance of AI technology or other algorithmic technologies.

- **Organizational Accountability**: All use of AI technology must be auditable and transparent, and be governed by a clear governance framework.

- **Transparency**: Where the Service uses AI technology that may have an impact on decisions that affect members of the public, the use of that technology must be made public to the greatest degree possible. Where full transparency may unduly endanger the

efficacy of investigative techniques or operations, the Service will endeavour to make publicly available as much information about the AI technology as possible, to assure the public of the reliability of the AI technology and the justifiability of its use. Where a decision assisted by AI technology may lead to the laying of criminal or other charges against an individual, the possible influence of the AI technology must be included in the disclosure provided to the Crown.

- **Privacy**: Use of AI technology must, to the greatest degree practicable, preserve the privacy of the individuals whose information it collects in line with 'privacy by design' principles.

- **Meaningful Engagement**: The adoption of specific AI technologies must be preceded by meaningful public engagement commensurate with the risks posed by the technology contemplated.

Purpose of Policy

The purpose of this Policy is to establish Board governance for the consideration of the use of new or enhanced technologies using AI, or of previously approved AI technology that is to be used for a novel purpose or in a novel circumstance, and to establish an assessment and accountability framework that addresses:

- The impact of the AI technology on the privacy, rights and dignity of individuals and communities, in accordance with the *Police Services Act* and its regulations (as well as successor legislation), Ontario's *Human Rights Code*, the *Canadian Charter of Rights and Freedoms*, and any other applicable legislation;
- The need for adoption new AI technologies to be done in a transparent manner, and contributes to equitable and effective policing services for all members of the public;
- Possible unintended consequences of the use of the AI technology in the provision of policing services in Toronto, prior to any adoption;
- A requirement for appropriate consultations to precede the procurement and deployment of new AI technologies that may have negative impacts on members of the public or the quality of policing services in Toronto;
- Mitigation strategies that seek to eliminate any identified unintended negative consequences stemming from the use of new AI technologies; and,
- A pre- and post-deployment, evidence-based evaluation and re-assessment of the AI technologies that are approved for procurement and/or use.

This Policy requires the thoughtful, evidence-based consideration of the benefits and risks of obtaining and deploying any new technology using AI, or novel uses of existing technologies, including impacts on public trust in the Service, community safety and sense of security, individual dignity, and equitable delivery of policing services. In particular, this Policy will ensure that decision-making examines and seeks to ensure that new technologies do not introduce or perpetuate biases to the greatest degree possible, including biases against vulnerable populations, including, but not limited to people with disabilities (physical and mental); children

and older persons; Indigenous, Black and racialized individuals; low-income individuals; and, members of LGBTQ2S+ communities.

Definitions

For the purpose of this Policy, the following definitions will apply:

AI Technology: goods and services, including but not limited to software and electronic devices, which collect information about members of the public or their actions, including personal information as defined under the *Municipal Freedom of Information and Protection of Privacy Act*, or make use of existing information about members of the public or their actions, and which use automated analytical problem-solving models to assist or replace Service Members in identifying, categorizing, prioritizing or otherwise making decisions pertaining to the information or the members of the public to which it pertains. AI technology includes, but is not limited to: machine learning technology, neural networks, natural language processing applications, predictive technologies, computer vision, and technologies which make predictions using algorithms trained on large data sets. Without limiting the foregoing, for the purpose of this Policy, "AI technology" will also include any goods or services whose procurement, deployment or use require that a privacy impact assessment be conducted in advance of its deployment or use.

New AI technology: any of: (1) AI technology never used before by the Service, (2) goods and services, including but not limited to software and electronic devices, already or previously employed by the Service which are enhanced through the application of AI in a manner that transforms the goods or services into an AI technology; (3) AI technology already or previously employed by the Service which is being considered for deployment for a novel purpose or in novel circumstances that may substantially change the data collected or used, including the content of the data, its granularity, and the purpose of data collection and use; (4) AI technology already or previously employed by the Service which is being enhanced through the use of new data that is substantially different from the data previously used, including the type of data, its granularity, or the manner in which it is obtained; and, (5) the linking of data from existing sources of information to create a new dataset for use by an AI technology.

Bias: systematically flawed output that is affected directly or indirectly by flaws in the design of the AI technology, training data, or the autonomous learning processes of the AI technology, to either misidentify certain types of subjects (individuals, objects, locations, etc.), or ascribe them with characteristics that disadvantage them based on illegitimate grounds (e.g., *Code*-protected grounds).

Data: any information collected and stored, whether locally or by a third party, which is used by the AI technology for the purpose of training, validation, testing, or generating output.

Biometrics: data on the measurements of physical and behavioural features of individuals (e.g., facial features, voice, gait) that could be used to identify the individual.

Human in the Loop: a process that ensures that any decisions or classifications made by the technology must be confirmed by a qualified human who can compare the input data with the output decision or classification, prior to any action taking place based on the output.

Explainability: AI technology is explainable when human users are able to comprehend the results created by the machine, why they were arrived at, and how changes to the input would have changed the outputs.

Training data: data provided to the AI technology for the purpose of enabling it to learn patterns and independently develop decision making algorithms.

Transactional data: data which is entered into a system which uses AI and that is used to generate output, but is not leveraged for training.

Policy of the Board

It is the policy of the Toronto Police Services Board that the Chief of Police:

Review and Assessment of New AI Technologies

1. Will develop, in consultation with the Information and Privacy Commissioner of Ontario, the Ministry of the Attorney General, the Anti-Racism Directorate, stakeholders, independent human rights experts, independent legal experts, independent technology experts, and affected communities, procedures and processes for the review and assessment of new AI technologies that will, at a minimum, establish:

    (a) That Service Members may not use new AI technologies prior to receiving approval and training in accordance with the procedure(s) and process(es);

    (b) That all Service Members must be trained to identify new AI technologies for the purpose of obtaining an approval in accordance with section 1(a);

    (c) Risk categories for new AI technologies based on their potential to cause harm, that include, at a minimum:

        i. Extreme Risk Technologies, which may not be considered for adoption, including:

            1. Any application where there is no qualified "human-in-the-loop". A qualified human must evaluate a recommendation from an AI tool before consequential action is taken, and be accountable for any decision made based on this recommendation;

2. Where use of the application results in mass surveillance defined as the monitoring of a population or a significant component of a population, or the analysis of indiscriminately collected data on a population or a significant component of a population;

3. Any application of AI in a life-safety situation, i.e., an application where the action of the AI technology could slow down the reaction time of the human operator, resulting in potential risk to life of members of the public or Service Members;

4. Any application that is known or is likely to cause harm or have an impact on an individual's rights, despite the use of mitigation techniques, due to bias or other flaws;

5. Any application used to predict or assign likelihood of an individual or group of individuals to offend or reoffend;

6. Any application making use of data collected in accordance with the Board's *Regulated Interaction with the Community and the Collection of Identifying Information* Policy, or any Historical Contact Data as defined in that Policy; or,

7. Where training or transactional data is known or thought to be illegally sourced, or where it is from an unknown source;

ii. High Risk Technologies, including:

1. Where training or transactional data is known or thought to be of poor quality, carry bias, or where the quality of such data is unknown;

2. Where training data can be influenced or biased by malicious actors;

3. Applications which link biometrics to personal information (e.g. facial recognition);

4. Where the proposed system could be used to assist in the identification of individuals for the purpose of their arrest, detention or questioning;

5. Where the process involved suggests an allocation of policing resources;

6. Where a system that otherwise merits a Moderate risk assessment lacks independent validation; or,

7. Where a system cannot be fully explainable in its behaviour;

iii. Moderate Risk Technologies, including:

1. Where the "human-in-the-loop" may have difficulty identifying bias or other decision failures of the AI; or,

2. Where training data is based on existing Service data;

iv. Low Risk Technologies, including any AI technology that both:

1. Does not fall under the categories of Extreme High Risk, High Risk, or Moderate Risk, and

2. Assists Members in identifying, categorizing, prioritizing or otherwise making administrative decisions pertaining to members of the public; and,

v. Minimal Risk Technologies, including any AI technology that does not fall under any of the preceding categories;

(d) The minimal risk analysis and privacy impact analysis that must be carried out for each level of risk in accordance with above subsection (c), as determined by an initial risk analysis, and the appropriate tools to carry out such impact analyses; and,

(e) The risk mitigation measures required for each level of risk (e.g., training, contingency planning);

2. Will make the procedures required under section 1, including a detailed risk assessment tool, available to the public on the Service's website;

Board Approval and Reporting Prior to Procurement, Utilization and Deployment

3. When contemplating procuring, utilizing or deploying new AI technology in its operations, will conduct a risk assessment of the AI technology, prior to the earlier of:

(a) Seeking funds for the new technology, including but not limited to applying for a grant, or accepting municipal, provincial or federal funds, or public or private in-kind or other donations;

(b) Acquiring the new technology, including acquiring such technology without the exchange of monies or other consideration;

(c) Using or deploying existing technology:

i. for a novel purpose;

ii. in novel circumstances, that may substantially change the data collected, including the content of the data, its granularity, and the purpose of data collection or use;

iii. for a purpose or in a manner not previously approved by the Board; or

iv. for a purpose or in a manner not practiced before the approval of this Policy; or,

(d) Entering into agreement to acquire, share, or otherwise use such technology;

4. Will not procure, utilize or deploy a new AI technology deemed to be of Extreme Risk;

5. Will not procure, utilize or deploy any new AI technology deemed to be of High or Moderate risk before reporting to the Board and obtaining its approval;

6. Will inform the Board, at the earliest possible opportunity, of the decision to procure, utilize or deploy a new AI technology deemed to be of low risk, and explain why the AI technology was ascribed this risk level;

7. When reporting to the Board in accordance with section 35, will describe, at a minimum:

(a) The operational need(s) the AI technology will address, including how use of the new AI technology will improve on current practices or operations;

(b) How the Service intends to use the AI technology;

(c) The risk level ascribed to the AI technology, why the AI technology was ascribed this risk level, and the rationale for continuing with the procurement, utilization or deployment requested despite the associated risk(s);

(d) The legislative authority for the collection of personal information;

(e) How the AI technology operates, including, where applicable, the source of the training data, what information will be collected, how and where information will be stored and how it will be disposed of, retention periods for the information collected, and evidence of the validity, accuracy and security of the AI technology under consideration, based on industry standards;

(f) The steps the Service will take or has taken to ensure the AI technology is used only in accordance with the *Police Services Act* and its regulations (as well as successor legislation), applicable privacy laws, Ontario's *Human Rights Code*, the *Charter of Rights and Freedoms* and other legislative and legal requirements, including training, and applicable governance;

(g) An evaluation of the AI technology's vendor, including its record with regard to data security and ethical practices;

(h) The results of any privacy impact and other assessment(s) that have been conducted;

(i) The feedback received from consultations with the Information and Privacy Commissioner of Ontario, the Ministry of the Attorney General, stakeholders and the general public, independent human rights experts, independent legal experts, independent technology experts, and affected communities;

(j) An analysis of possible unintended consequences of the proposed use of the AI technology from legal and human rights perspectives, including the potential for disproportionate impacts on *Human Rights Code*-protected groups, and steps the Service will take to mitigate these unintended consequences;

(k) Where applicable, a legal analysis of potential challenges to the admissibility of evidence generated or impacted by the AI technology in criminal proceedings;

(l) The findings of any risk analyses carried out in accordance with section 1(d) above, and any additional analysis as appropriate, including any analyses required by the Information and Privacy Commissioner of Ontario;

(m) Any reports and documentation used in the evaluation of AI technology;

(n) A mitigation plan to:

i. Mitigate the risks posed by the implementation of the AI technology, including risks of biased policing, infringement of privacy or other rights, chilling effects on freedom of expression, and risks of abuse or unauthorized access to information, and including the mitigation of any bias or quality issues in the training data used by the AI technology;

ii. Ensure that any use of the AI technology will be audited to ensure adequate and lawful use, in accordance with the purposes approved by the Board, and to monitor errors; and,

iii. Notify the Information and Privacy Commissioner of Ontario and/or any other applicable legal authority of any significant privacy breaches or other significant malfunctions that may result in harm to individuals, communities or Service Members, or may impact criminal proceedings;

(o) The estimated cost of acquiring and implementing the AI technology, including the cost of adequate training in the use of the AI technology, and any additional costs or savings expected from the implementation of the AI technology; and,

(p) Proposed indicators that will be tracked by the Chief of Police aimed at determining whether the AI technology is achieving its intended goal and

whether its deployment has had any unintended consequences, until it is determined by the Board that monitoring is no longer required;

8. Will develop and implement a public engagement strategy, commensurate with the risk level assigned to the new AI technology, to transparently inform the public of the use of the new AI technology that collects data about members of the public or assists Service Members in identifying, categorizing, prioritizing or otherwise making decisions pertaining to members of the public, prior to its deployment; and,

9. Will develop and implement a strategy to communicate to the Crown the risks of an AI technology that require judicial authorization for its application, or which may impact any criminal proceedings.

It is further the policy of the Board that:

10. The Board will review the reports submitted in accordance with section 5 and may:

    (a) Request or solicit an independent review of the recommendations made by the Chief;

    (b) Determine that additional analysis is required prior to approval of the procurement, deployment or use of the new AI technology;

    (c) Determine that the Service may initiate a pilot process for the use of the new AI technology to better assess it, and identify the parameters of the pilot in a manner that mitigates any risks of biased decision-making by Service Members; or,

    (d) Determine that the Service may initiate the procurement, deployment or use of the new AI technology, and identify any additional analysis, monitoring, auditing and reporting requirements beyond the ones required by this Policy that are to be imposed once use of the AI technology commences.

Monitoring and Reporting

It is the policy of the Board that the Chief of Police:

11. Will monitor the indicators approved by the Board under Section 7(p), from the initiation of deployment and until 12 months after full deployment of new AI technology deemed to be of Moderate risk, or until 24 months after full deployment of new AI technology deemed to be of High risk;

12. Will report to the Board, within 15 months of full deployment of a new AI technology deemed to be of High or Moderate risk, and again within 27 months of full deployment of a new AI technology deemed to be of high risk, with such reporting describing :

(a) How the AI technology has generally been deployed or utilized within the first period until 12 (or 24) months from full deployment, including with respect to compliance with applicable privacy laws and other legislative and legal requirements;

(b) The performance as measured by the indicators approved by the Board under Section 7(p) of this Policy;

(c) What concerns the Chief of Police has seen raised by members of the public or Service Members, and how the Chief has acted to address those concerns where appropriate;

(d) For AI technology deemed to be of High risk, the results of a post-deployment public consultation on the impacts of the deployment;

(e) Whether the Chief intends to continue using the AI technology in the same manner or in a different manner in the future; and,

(f) Where the Chief intends to continue using the AI technology, the key performance indicators that the Chief will continue to monitor indefinitely to ensure the continued quality of the AI technology's performance, and that no new unintended consequences emerge through its use; and,

It is also the policy of the Board that:

13. The Executive Director shall create a method for members of the public to submit concerns pertaining to specific AI technologies used by the Service through the Board's website, and

    (a) Where concerns are expressed with regard to an AI technology deemed to be of Moderate or High risk, for which the Service has not yet submitted the report required by section 12, will append a summary of the concerns to the report when it is brought before the Board; or

    (b) Where concerns are expressed with regards to an AI technology for which the Service has already submitted the report(s) required by section 12, or with regards to an AI technology deemed to be of Low or Minimal risk, will:

        i. if the Executive Director determines that the concern raised likely demonstrates that an AI technology was erroneously assessed at a lower risk level than appropriate in accordance with section 1(c), will report on the nature of the concern to the Board at the earliest possible opportunity; and,

        ii. otherwise, report annually to the Board with a summary of the concerns raised by members of the public; and

(c) Where a communication from a member of the public amounts to a complaint under Part V of the *Act* or successor legislation, will advise the individual or their right to file a complaint with the Office of the Independent Police Review Director (or successor entity), or forward the communication to the Chief of Police, as appropriate, and inform the complainant of this action;

14. The Board will review the reports provided in accordance with above section 12, and determine whether the Service may continue to use the AI technology in question, and whether any additional analysis, monitoring, auditing and reporting requirements are to be imposed, and in particular whether the Chief of Police must continue to monitor the indicators approved by the Board under Section 7(p); and,

15. All reports required by this Policy will be considered by the Board in its regular public meetings, with the exception of any information provided in the report for which confidentiality is maintained in accordance with applicable law, in which case only that information will be provided to the Board separately as a confidential attachment to the public report.

Continuous Review

It is also the policy of the Board that the Chief of Police:

16. Will initiate immediately a process to identify and conduct a risk analysis of all AI technologies currently in use by the Service, to be completed no later than December 2024, and report to the Board upon its completion with a summary of its findings;

17. Will post immediately on the Service's website, and maintain up to date with the most accurate available information, a list of all AI technologies currently in use by the Service that are deemed to be of High, Moderate or Low risk, including the following information:

    (a) For AI technologies deemed to be of High or Moderate risk:

        i. Name and manufacturer/developer,

        ii. Purpose of the technology,

        iii. How the technology is used by the Service,

        iv. What information is collected by the technology,

        v. What persons or under what circumstances can the technology be expected to be used, and,

        vi. All reports submitted by the Chief to the Board with regards to the AI technology, as required under this Policy or subsequent Board decisions;

    (b) For AI technologies deemed to be of Low risk:

i. Name and manufacturer/developer, and

ii. A brief description of the type of technology (e.g., speech-to-text);

18. Will terminate the use, immediately upon identification, of any AI technology in use by the Service prior to the adoption of this Policy, which is deemed to be of Extreme risk, and inform the Board of this action with a description of the AI technology that was identified, the reason that it was deemed to be of Extreme risk, and an assessment of potential harms that were caused to individuals, communities or Service Members, and possible impacts on criminal proceedings, as a result of its use;

19. Will report to the Board, as soon as it is identified, concerning any AI technology in use by the Service prior to the adoption of this Policy, which is deemed to be of High or Moderate risk, including:

    (a) the reason that the AI technology was deemed to be of this risk level, and,

    (b) a plan to:

        i. pause the use of the AI technology within no longer than three months,

        ii. evaluate the risk and any potential harms resulting from the use of the AI technology,

        iii. develop a mitigation plan, and

        iv. seek the approval of the Board for the continued use of this AI technology;

20. Will review at least once every two years in the case of an AI technology deemed to be of High risk, and at least once every five years in the case of AI technology deemed to be of Moderate risk, the continued use of any AI technology based on:

    (a) the quality of the AI technology, its outputs, and associated key performance indicators; and,

    (b) the continued need for the use of the AI technology; and;

21. Will review at least once every five years the use of any AI technology deemed to be of High, Moderate or Low risk to ensure that the AI technology has not been put to use for a novel purpose or in novel circumstances that may substantially change the data collected or used, in a manner that would constitute a new AI technology, or the risk level of the AI technology, and, where it is found that an AI technology has been put to a

new use in this manner, will report to the Board as soon as possible, in accordance with section 4.

It is also the policy of the Board that:

22. The Board will review the Policy at least once every three years to ensure that the Policy successfully achieves its identified purpose. In particular, the Board will review any instance where a report was made in accordance of section 13(b)i, to consider whether any changes are required to minimize the potential of misclassifications of risk.