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	
DERIVATION		

Guiding Principles

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. 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.

Under s. 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. As such, it is important for the Board and Chief to engage in a constant dialogue about how technology will be procured, implemented and used in the provision of policing in Toronto.

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 legislation fully regulates the use of AI technologies, and the Province has not yet developed comprehensive guidelines for the use of such technologies in policing. As a result, the Board must consider new technologies utilizing AI that the Service intends to use in the provision of policing in Toronto, and establish policies with regard to the use of such technologies, as needed. In its review, 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 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.

Purpose of Policy

The purpose of this Policy is to ensure that the Board, as the oversight body representing the public interest in providing policy direction to the Chief of Police, is involved in the consideration of the use of new or enhanced technologies using AI, or of previously approved technology that is to be used for a novel purpose or in a novel circumstance, and to establish a protocol for reporting to:

- Preserve the privacy, rights and dignity of individuals and communities, in accordance with the *Ontario Human Rights Code*, the *Canadian Charter of Rights and Freedoms*, and any other applicable legislation;
- Ensure that the adoption of new technologies is done in a transparent and trustworthy manner, and results in the advancement of equitable and effective policing services for all members of the public;
- Ensure that consideration is given to possible unintended consequences of the use prior to the adoption of new technologies in the provision of policing services in Toronto;
- Ensure that appropriate consultations precede the procurement and deployment of new technologies that may have negative impacts on members of the public or the quality of policing services in Toronto;
- Develop mitigation strategies to eliminate any unintended negative consequences from the use of new technologies; and
- Ensure that a pre- and post-deployment evidence-based evaluation and re-assessment takes place.

This Policy will ensure the thoughtful 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, it will help to ensure that new technologies do not introduce or perpetuate biases, including biases against vulnerable populations such as people with disabilities, children and older persons, Indigenous, Black and racialized communities, low-income and LGBTQ2+ communities, into policing decisions.

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, and computer vision. 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 their 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: consistently flawed output that is affected by flaws in the design of the AI technology or training data, to either misidentify certain types of subjects, 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 human who can compare the input data with the output decision or classification, prior to any action taking place based on the output.

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 experts and stakeholders, 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 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 that include, at a minimum:

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

            1. Any application where there is no "human-in-the-loop". A human must evaluate a recommendation from an AI tool before consequential action is taken;

            2. Where use of the application results in mass surveillance defined as the indiscriminate covert monitoring of 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 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; or

5. 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 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); or

4. 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 the process involved suggests an allocation of resources;

   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 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 minimum 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 harm 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. Will not procure, utilize or deploy a new AI technology deemed to be of extreme risk;

4. When contemplating procuring, utilizing or deploying new AI technology in the field, will conduct a risk assessment of the AI technology and report to the Board where the AI technology is found to be of high or moderate risk, 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;

5. When reporting to the Board in accordance with section 04, 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;

   (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, what information will be collected, how information will be stored and how it will be disposed of, and evidence of the validity and accuracy of the AI technology under consideration;

(f) The steps the Service will take or has taken to ensure the AI technology is used only in accordance with applicable privacy laws, the *Human Rights Code* and the *Charter of Rights and Freedoms* and other legislative and legal requirements, including training, and governance structures;

(g) The results of any privacy impact and other assessment(s) that have been conducted, and consultations with the Information and Privacy Commissioner of Ontario, the Ministry of the Attorney General and other stakeholders, independent human rights, legal and technology experts and affected communities, as appropriate in light of the potential risks posed by the contemplated technology;

(h) An analysis of possible unintended consequences of the proposed use of the AI technology, including possible effects on procedural fairness, due process, gender and race equality, or disproportionate impacts on *Human Rights Code* protected groups, and steps the Service will take to mitigate these unintended consequences;

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

(j) 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;

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

(l) A mitigation plan to mitigate the risks posed by the implementation of the AI technology;

(m) The estimated cost of acquiring and implementing the AI technology, and any additional costs or savings expected from the implementation of the AI technology; and,

(n) Proposed indicators that will be tracked by the Chief of Police until at least 12 months after full deployment of the new AI technology to determine whether the AI technology is achieving its intended goal and whether its deployment has had any unintended consequences;

6. Will not procure, utilize or deploy any new AI technology deemed to be of high or medium risk before obtaining the Board's approval;

7. Will inform the Board, at the earliest possible time, 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; and

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.

It is further the policy of the Board that:

9. The Board will review the reports submitted in accordance with section 5 and determine whether the Service may initiate the procurement, deployment or use of the new AI technology, and whether any additional analysis, monitoring, auditing and reporting requirements beyond the ones required by this Policy are to be imposed.

Monitoring and Reporting

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

10. Will monitor from the initiation of deployment and until 12 months after full deployment of the new AI technology deemed to be of high or medium risk the indicators approved by the Board under Section 5(n);

11. Will report to the Board, within 15 months of full deployment of a new AI technology deemed to be of high or medium risk, with such reporting describing :

    (a) How the AI technology has generally been deployed or utilized within the first period until 12 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 5(n) 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;

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

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

12. Will continue to track the indicators approved by the Board under section 5(m)5(n) until it is determined by the Board that no additional monitoring is required.

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 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 Medium or High risk, for which the Service has not yet submitted the report required by section 11, 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 required by section 11, or with regards to an AI technology deemed to be of Low or Minimal risk, will:

        i. If the Executive Director finds that the concern raised likely demonstrates that an AI technology was erroneously assessed as of 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, will advise the individual or their right to file a complaint with the Office of the Independent Police Review Director or successor role, or forward the communication to the Chief of Police, as appropriate, and inform the complainant of this action; and

14. The Board will review the reports provided in accordance with above section 11 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.

Continuous Review

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

15. Will post on the Service's website no later than December 2024, and maintain up to date, a list of all AI technologies currently in use by the Service that are deemed to be of High, Medium or Low risk, including the following information:

    (a) For AI technologies deemed to be of high or medium 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, and

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

(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);

16. Will terminate the use, immediately upon identification, and no later than December 2024, 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 and the reason that it was deemed to be of Extreme risk;

17. Will report to the Board, as soon as it is identified and no later than December 2024, of any AI technology in use by the Service prior to the adoption of this Policy, which is deemed to be of High or Medium risk, including:

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

(b) a plan to evaluate the risk and any potential harms resulting from the use of the AI technology, develop a mitigation plan, and seek the approval of the Board for the continued use of this AI technology;

18. Will review at least once every five years the continued use of any AI technology deemed to be of High or Medium risk 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

19. Will review at least once every five years the use of any AI technology deemed to be of High, Medium 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 5.