Community Safety Knowledge Alliance
Research to Practice to Alignment

AI and the Future of Canadian Policing: Implications and Strategic Considerations

Comprehensive Report for Sponsoring Chiefs of Police

October 2026

Community Safety Knowledge Alliance
cskacanada.ca

This page intentionally left blank

Executive Summary

Artificial intelligence is reshaping the nature of work, education, government, and the public realm. It has already begun to influence Canadian policing through investigative analytics, automation, digital evidence management, and emerging decision-support tools. AI is also a threat multiplier: criminal networks, violent ideological actors, and foreign adversaries are using AI to scale fraud, disinformation, recruitment, and cyber operations. The strategic question for police leadership is not whether AI will appear in policing, but whether its adoption will be governed in ways that strengthen operational effectiveness and preserve public trust, legal legitimacy, and institutional resilience - and whether its use by malicious actors will be understood and tackled effectively.

This paper was prepared for Canadian police leaders - chiefs, police boards, and policymakers - who, together, are responsible for operational effectiveness, institutional legitimacy, public trust, and long-term organizational resilience. This report is intended to enable careful reflection and decisions about the set of important emerging challenges related to AI that are on the minds of the communities they serve. Decisions about AI adoption will either strengthen or undermine the public trust and legitimacy that are foundations to democratic policing in our country. The research team comprised researchers and practitioner leaders with extensive expertise in operational policing, police leadership, governance, community safety, and law.

Drawing on a synthesis of scholarly and grey literature, exploratory interviews with subject matter experts conducted between February and May 2026, and a structured review of the draft by a multidisciplinary expert panel in September 2026, this report assesses AI's emerging impact on policing, identifies conditions under which its use may undermine legal obligations, ethical commitments, and public confidence, and offers practical guideposts for police leaders navigating this rapidly evolving terrain.

Our key purpose is to support well informed, principled decision-making. This paper reflects conditions as of June 2026 and was revised in October 2026 to incorporate feedback from an multidisciplinary panel of experts. Given the pace of change, its guidance should be revisited at least annually.

Key Findings

The following five areas capture the most significant ways AI is affecting Canadian policing today and in the near term. Each presents both genuine opportunity and significant risk. Furthermore, each carries governance implications that police leaders and boards cannot treat as purely technical or operational matters.

1. AI Offers Real Operational Benefits - Under the Right Conditions

AI's most immediate value to policing lies in its capacity to manage scale, complexity, and speed. Systems can automate routine tasks (e.g., interview transcription, report drafting, call triage subject to safeguards), enhance investigative capacity through pattern detection, advanced analytics and image matching across large datasets, and support resource deployment and situational awareness.

These benefits are genuine. But they are contingent. AI adds value where it augments professional judgement and addresses a clearly defined problem, not where it displaces accountability or arrives as a solution in search of a problem. Administrative uses and operational uses that directly affect the public also carry different risk profiles and should not be governed as though they were the same. Poorly designed or poorly governed systems introduce risks that can undermine the very objectives they are meant to serve, including reliability failures, evidentiary complications, and erosion of the human discretion that policing requires.

2. The Regulatory Environment Is Evolving and Demanding

Police services deploying AI must navigate an interconnected legal and policy framework spanning constitutional law (e.g., the *Charter of Rights and Freedoms*), federal and provincial privacy legislation, criminal procedure and evidentiary rules, administrative law, and digital governance regimes. In June 2026 the federal government introduced Bill C-36, the *Protecting Privacy and Consumer Data Act* (PPCDA). If enacted, this legislation will repeal Part 1 of *Personal Information Protection and Electronic Documents Act* to establish a modern federal private-sector privacy system for Canada. Bill C-36 aims to strengthen consent-based collection of personal information, and would create a new federal regulator, the **Digital Safety and Data Protection Commission of Canada** to enforce online safety laws, oversee social media platforms and AI chatbots, and protect children from harmful digital content. The Commission would have expanded enforcement powers, including the ability to impose significant administrative monetary penalties.

Canadian courts are also signalling closer scrutiny of AI in proceedings, and disclosure obligations are likely to extend to algorithmic processes. Legislation, procurement rules, and vendor auditing practices are all struggling to keep pace with the technology itself. AI adoption is never simply a procurement decision. It carries constitutional, privacy, evidentiary, and cybersecurity governance dimensions that require structured review before any deployment, and governance that adapts as tools, vendors, and the law change.

3. Ethical and Practical Risks Are Significant

AI systems trained on historical data risk reproducing and amplifying past patterns of bias, over-policing, and systemic discrimination, particularly affecting Indigenous, racialized, and marginalized communities. The Clearview AI controversy illustrated that legitimacy failures arise not only from technical errors but from transparency gaps and unauthorized use.

Many AI systems operate opaquely, creating accountability challenges in court and undermining public confidence. Where proprietary “black box” systems generate or shape evidence, police and the Crown may be unable to meet disclosure obligations or defend that evidence when it is challenged in court. Vendor performance claims are difficult for lay persons to verify, and independent empirical validation of many police AI tools remains limited. Human review is no guarantee on its own: automation bias can lead reviewers to defer to confident-looking outputs. Public skepticism is substantial: recent Angus Reid polling found that 68% of Canadians believe government needs to heavily regulate AI even if it slows development, and most doubt any government can keep pace. Ensuring that affected individuals can understand and meaningfully

challenge AI-influenced decisions engages Charter and disclosure obligations and is a precondition for sustained institutional legitimacy.

4. AI Introduces New Institutional and Security Risks

AI concentrates sensitive data, expands cyber-attack surfaces, and generates new categories of operational and strategic risk. As AI tools become embedded in workflows, police agencies risk vendor lock-in and platform dependence, potentially losing independent audit capacity and operational control.

“Shadow AI” (personnel using unapproved AI tools or personal devices/accounts) represents an under-recognized compliance and security risk. Foreign ownership concerns (particularly regarding platforms subject to Chinese or US regulatory reach) and data sovereignty require active attention. Perhaps most insidiously, “decision-shaping” can occur when AI platforms begin to determine what is visible, prioritized, and acted upon, gradually rewiring organizational behaviour in ways that are difficult to detect and reverse.

5. AI-Enabled Threats Are Escalating

Criminal, ideological, and state actors are already exploiting AI to accelerate fraud, disinformation, child sexual exploitation, and hybrid threats. Deepfakes, automated fraud, synthetic media, and AI-enabled radicalization will outpace traditional investigative approaches unless adaptive countermeasures are developed at the speed of cyber.

Police services must also contend with insider threats linked to AI-enabled online radicalization of both civilian and sworn members. Responding effectively requires integrated intelligence, close coordination with cybersecurity and national security partners, positive community relations, and sustained leadership attention to workforce integrity and organizational culture.

The Strategic Imperative

Some AI adoption in Canadian policing is unavoidable. With organized crime and others already leveraging these tools, police cannot afford to disengage. But AI adoption is not a neutral modernization step. AI systems do not merely improve technical efficiency - they reshape what is visible, prioritized, and acted upon. The strategic issue for police leaders and boards is to ensure that AI strengthens effectiveness without eroding fundamental rights and institutional legitimacy. This means treating AI as a governed capability: embedded within professional judgement, constrained by law, and accountable to the communities it serves.

Without proactive governance, agencies risk:

- gradual erosion of trust and legitimacy inside and outside the organization;
- loss of critical thinking and professional judgement;
- disproportionate impact on racialized and equity-deserving communities;
- normalized surveillance practices;
- evidentiary risks in court; and
- vendor dependence and data sovereignty failures.

These harms tend to accumulate, making them difficult to reverse. AI governance is not a communications strategy - it is institutional infrastructure. Successful AI adoption depends as much on people, organizational culture, and community relationships as it does on technology.

Summary of Recommendations

The following summarizes the paper's 21 recommendations, grouped by theme. Five are designated as a Year-1 Core: foundational controls every service should put in place first, regardless of size. The remainder can be phased in as capacity matures. Full detail appears in the Recommendations section.

Governance and Oversight

1. Designate AI adoption as a standing item on the executive governance agenda, with a named executive lead, an inventory of AI tools in use, and quarterly board reporting on deployments, incidents, and review findings. (Year-1 Core)
2. Establish a standing external review panel comprising technical, human rights/privacy, and community expertise, with a mandate to review high-risk deployments annually, report findings publicly, and access method-level detail confidentially.
3. Conduct annual legal review of investigation protocols against current case law on AI-related evidence integrity and Crown disclosure obligations.
4. Require annual third-party performance, bias and social compliance[1] audits for all AI vendor contracts, including re-validation after material updates, with results delivered to the board.

Decision Gates

- Adopt a tiered risk-classification scheme that distinguishes administrative from operational uses, requires board approval for high-risk deployments, and triggers re-assessment whenever a tool's approved scope changes. This classification scheme should be reviewed and updated as required. (Year-1 Core)
- Require a documented joint sign-off from legal counsel, privacy officer, and IT security before any AI tool deployment, extended to municipal IT and procurement partners where services are shared.
- Identify or help establish a provincial or national vetted-tools registry within 12 months.
- Define (in advance) the conditions that require a tool to be rejected, restricted, or withdrawn (the "zero-option"), and publish a register of tools in use and tools withdrawn.

Procurement and Risk Management

[1] A social compliance audit involves an assessment of whether an organization and its suppliers adhere to laws, standards and ethical practices regarding labour conditions, human rights, worker health and safety, etc.

- Require a documented problem statement, including a 'safety for whom?' assessment, before any AI acquisition, and score cybersecurity, inspectability, explainability, data sovereignty and local control in RFPs together with cost and performance.
- Develop standard contract templates requiring data-ownership terms, minimum cybersecurity standards, disclosure sufficient for court proceedings, vendor expert testimony, notice and re-validation before updates, change-of-ownership protections, cost controls, and exit/data-portability clauses. (Year-1 Core)
- Update acceptable-use policies within 90 days to explicitly prohibit unapproved AI tools (e.g., personal ChatGPT or Gemini accounts) as shadow AI, and identify approved alternatives. (Year-1 Core)
- Require comprehensive knowledge transfer provisions in all technology contracts to support in-house capability development.
- Formalize service level agreements and a cross-governance framework with municipal IT and procurement bodies, so that police retain control over Charter, privacy, and evidentiary compliance in shared-service environments.

Workforce and Organizational Preparedness

- Deliver mandatory AI-literacy training to all employees within 12 months, refreshed annually, or as needed, backed up with simple checks so staff don’t rely on AI results without verifying them. (Year-1 Core)
- Add online radicalization indicators to insider threat screening, recruit vetting, and periodic review, in consultation with unions and associations.
- Sequence high-risk AI deployments to internal capacity: deploy no new iteration of a high-risk tool until the prior one has completed its first independent review. Lower-risk administrative tools may proceed on a faster track.

Transparency and Accountability

- Produce a plain-language explanation template for each deployed AI tool, covering what it does, what data it uses, and how its outputs can be challenged.

Community Engagement

- Treat structured community consultation for medium- and high-risk tools as a non-waivable requirement, with meaningful decision weight for the communities most likely to be affected and published responses to their input.
- Allocate a specific budget line and executive lead for community engagement within the AI governance framework, with annual adequacy review.

Fragmented Capacity

- Advocate for a shared regional or provincial AI evaluation body at the next relevant chiefs’ association meeting, with a working group established within six months and access to shared technical, legal, and privacy expertise for smaller services.

- Pilot a human rights AI impact assessment framework as part of the governance architecture, with a view to making it standard for high-risk tools.

Glossary

Agentic AI: AI systems that can independently plan and take actions to achieve a goal, rather than simply responding to a user's request. They can perform multi-step tasks, use tools, and make decisions with limited human intervention. Examples include: analyzing large volumes of information, coordinating investigative workflows, or managing administrative processes.

Algorithm: A step-by-step set of instructions that tells a computer how to solve a problem or make a prediction.

Algorithmic bias: When a computer system produces unfair or skewed results, often because the data it was trained on reflects existing inequalities, gaps, or patterns of discrimination.

Artificial intelligence (AI): Computer systems designed to perform tasks that normally require human thinking, such as recognizing patterns, understanding language, making predictions, or helping with decisions.

AI governance: The set of rules, processes, oversight structures, and accountability mechanisms that guide how AI systems are designed, acquired, used, monitored, and reviewed. AI governance ensures that AI is deployed in ways that are lawful, ethical, transparent, secure, and aligned with organizational values and public expectations, not just technically effective.

Automation bias: The tendency to over-trust automated outputs, accepting a system's recommendation, match, or summary without sufficient independent scrutiny, particularly under time pressure.

Data sovereignty: The principle that individuals, communities, organizations, or nations have the right to govern how data about them is collected, used, stored, shared and protected. It is based on the idea that those who are the subject of the data should have a meaningful role in determining who can access it, how it is used and for what purposes.

Function creep: The gradual expansion of an AI tool beyond its originally approved purpose, scope, or data sources, often without renewed review or approval. Also known as scope creep.

Generative AI (GenAI): A type of AI that can create new content, such as text, images, audio, video, or code, in response to a query (a "prompt"). It is based on patterns it has learned from large amounts of data that it has been trained on. Instead of simply analyzing information, generative AI produces new material based on what it predicts should be the best response to a query.

Shadow AI: The use by personnel of AI tools that have not been approved by the organization, such as personal generative AI accounts. Shadow AI is a growing subset of "shadow IT."

Zero-option: The retained authority, supported by predefined conditions, to decline, restrict, or withdraw an AI tool or system at any stage, regardless of prior investment or vendor commitments.

Table of Contents

Executive Summary 3
Glossary 9
Introduction 11
Methodology 14
AI Use in Policing: Impacts and Implications 16
1. Enhancing Effectiveness and Efficiency in Policing 16
2. Broader Regulatory Context 18
3. Legal, Ethical and Practical Implications 20
4. Threats, Risk and Security 30
5. Criminal, Ideological, and State Actors 36
Conclusions 39
Recommendations 41
Impact of Recommendations 45
Areas Warranting Further Study 46
References 48
Acknowledgements 56
Report Authors 57
About CSKA 58
AI Disclosure 58
Appendix A: Sample Decision Guideposts for AI Adoption 59

Introduction

> *"If you like some scenarios better than others, you can be aware of the ones you don't like and look for signs of them, and also look at signs of the ones you want to have come to pass, and lean differentially toward them. That's how you negotiate your way into a future you were glad of. It's done incrementally by, among other things, lots of individuals and some institutions, and that's how we grapple our way, muddle our way forward."*
>
> - *Stewart Brand[2], futurist and writer*

This paper is expressly written for Canadian police leaders - chiefs and their executive teams, police boards, and policymakers - who, together, are responsible for operational effectiveness, institutional legitimacy, public trust, and long-term organizational resilience. It synthesizes the implications of recent analyses on artificial intelligence (AI), policing, and community safety and well-being, with particular attention to governance, risk, and societal impact.

Given the rapid pace of technological and regulatory change, readers should note that this paper reflects conditions as of June 2026, and was revised in October 2026 to reflect the feedback of an multidisciplinary expert review panel. While we have made every effort to ensure accuracy, some information may have evolved by the time of reading. For that reason, our guidance is principles-based, so that it remains useful as specific tools, vendors, and laws change.

Our purpose is to assess AI's emerging impact on policing, including its potential benefits and risks to legal compliance, public trust, and community safety - and to offer practical guidance for responsible adoption that is effective, appropriate and trust building.

AI-based technologies, which mainly involve the use of digital algorithms to support data collection, data analysis and automation, are not only changing the nature of the tools that police use, but also the conditions in which policing and community safety and well-being (CSWB) are produced: types of criminal activity; the stability of local economies and employment; household security; information integrity (perceptions of crime and safety and police legitimacy within a community population, as well as the informational bases for partnership and collaboration); and the pace at which AI disruptions and other shocks may spread through communities and community organizations.

The term artificial intelligence attracts as much confusion about what it means, as it gets increasing attention in the media. To help clarify the situation, Bender and Hanna (2025) described AI in terms of the five main types of work that these systems automate:

- **Decision making:** Systems used to automate decisions, such as the use of predictive algorithms used to screen resumés or in bail determination processes.

[2] Quoted in Rose (2026)

- **Classification:** The automation of categorization activities, such as pattern matching, comparing static images to an existing database of photos, or person identification based on video surveillance.
- **Recommendation:** The presentation of information based on previous behaviour, such as search history or other uses of a system.
- **Transcription and translation:** This involves the automatic translation of data from one form to another form, such as speech recognition or speech-to-text software. It also includes automated licence plate readers (ALPRs).
- **Text and image generation:** These are synthetic media systems that generate images or text based on written or spoken prompts that specify what the desired output is to be. Common generative AI technologies include Anthropic's Claude and OpenAI's ChatGPT and DALL-E. Image generation technologies used for malicious purposes are commonly known as deepfake systems. When used to generate realistic nude images of targeted individuals, these systems are sometimes known as "nudification" technologies or "nudify" applications.

Amid the swirl of definitions, expectations and concerns, many suggest that AI is poised to have a significant impact on society, reshaping people's work, education, government, the public realm, relationships and even how individuals think about themselves.[3] Those who are convinced of the transformative power of AI tend to fall into one of two perspectives. Advocates of these viewpoints are often known as either boosters or doomers.[4]

> **Boosters** are those who assert that AI is an unstoppable force for net good, and that any efforts to slow the pace of AI adoption would be unrealistic and misguided. This is often accompanied by calls to avoid missing out on the opportunities that AI adoption might entail.
>
> **Doomers** frame AI as a potentially unstoppable threat to the survival of humanity, if it is not constrained right now. This second group includes such prominent voices as Canadian Nobel Laureate, Geoffrey Hinton, who has called for a slowing or even a moratorium on AI development until more robust guardrails can be installed[5].

Both AI boosters and doomers see AI dominance over many aspects of human life as an inevitability, and try to create a sense of urgency around responding to the issues they raise. Most call for AI development to be accompanied by various protocols and guardrails under the heading of "AI safety". These guardrails may be technical, legislated or anchored in corporate policies and culture.

In contrast, AI skeptics (e.g., Bender & Hanna, 2025; Doctorow, 2026) question the assumption that the rapid evolution and implementation of AI technologies cannot be resisted. They assert the importance of prioritizing human connections and processes, and advocate for human

---

[3] Carleton & Guirado (2026); Wischmeyer (2019)
[4] e.g., Bender & Hanna (2025)
[5] Cecco (2026)

choice and a greater degree of control over AI technologies and the business models that underlie their design, development and implementation. Unlike boosters and doomers, AI skeptics advocate for a less hurried engagement with AI. In some cases, they may call for varying degrees of strategic refusal around the adoption of AI technologies.

A similarly guarded viewpoint considers the human rights dimensions of digital data (e.g., Wong, 2023). This perspective re-positions people from being the subjects of digital technologies to becoming stakeholders of digital data and technologies, based on the values of autonomy, dignity, equality and social connection.[6]

The February 2026 shooting in Tumbler Ridge, BC, which left eight dead, brought into stark focus one of the ways that AI misuse can result in great harm. According to media reports, OpenAI staff had internally flagged the alleged shooter's ChatGPT conversations about gun violence in the months prior to the attack.[7] Reporting indicated that, despite those warnings, OpenAI leadership overruled its safety team and neither government officials nor local police were notified. Canadian politicians offered differing responses: one called for stricter internal standards by technology companies; another called for federal notification requirements when AI firms identify credible warning signs.[8] Canada currently has no mandatory reporting requirements for platforms, and the threshold for triggering such reporting remains undefined. This intersection of public safety and individual privacy issues illustrates how gaps in policies, practices and regulation, along with diverging approaches to responsible AI use, may contribute to elevated vulnerabilities.

Mainstream economic institutions anticipate that AI will affect at least one in five jobs, through substitution, task redesign, and wage pressure, and that there will be uneven impacts across industrial and service sectors and worker groups.[9] In Canada, the proportion of affected employees could be as high as 60%.[10] That unevenness matters for policing because it can translate into localized spikes in perceived insecurity, opportunistic crime, potential increases in the demand for health and social services and possibly also, social unrest.

The AI transition is also arriving amid a broader "polycrisis"[11] of: climate impacts such as extreme weather events; national unity issues; geopolitical tensions and conflicts, including hybrid operations by Canada's adversaries; food insecurity; economic volatility and a contested information environment. These factors may interact in ways that can strain services, community cohesion as well as trust in the capacity and willingness of public institutions to safeguard the communities they serve. Faced with these conditions, and under the demands of austerity and economic reorganization, police agencies are likely to experience growing pressure to conduct their work in more efficient ways.

---

[6] Wong (2023)
[7] Globe & Mail Editorial Board (2026)
[8] Ibid
[9] e.g., World Economic Forum (2025)
[10] Mehdi & Morissette (2024)
[11] Lawrence, et al. (2024)

The value proposition promoted by the providers of most AI-based services and equipment often turns on claims that these platforms and tools will confer greater cost savings and improve the tempo of consequential operations. This can be compelling where police agencies are faced with labour shortages, growing stakeholder expectations for efficiency and the challenges of managing complex operations. The interest of police in adopting compelling new tools is also an element of their practice cultures that cannot be under-estimated. This makes the starting point important. Adoption should begin with a clearly defined operational or community safety problem, and the question of "safety for whom?", rather than with a vendor's proposal.

The purpose of the present review is to help police leaders, and those involved in governance, to better understand and contend with the ways that AI technologies may impact Canadian policing in the years ahead.

Objectives

Our objectives were threefold:

1. to describe and illustrate how AI is already influencing the environment and practices of policing;
2. to assess where it can offer genuine value, and to surface the conditions under which its use may undermine legal obligations, ethical commitments, organizational effectiveness, public confidence and outcomes related to CSWB; and
3. to offer a set of guideposts for police agencies that are attempting to navigate this emerging AI-infused world.

Methodology

This review draws on three complementary sources of evidence: a synthesis of scholarly and grey literature, a series of interviews with subject matter experts, and a multidisciplinary expert panel review of the draft report.

Literature Synthesis

This study employed an iterative, narrative review of peer-reviewed scholarly literature and grey literature (published and unpublished reports, investigative journalism, expert opinions, and issues papers).

Initial discovery-focused searches of purposively selected search engines and AI platforms (Google Scholar, Perplexity) and the internet used predefined keywords, which were progressively refined as new concepts and key authors emerged. Snowball searching[12] of reference lists and additional cited sources was used to follow promising lines of inquiry and reach saturation of recurrent themes. All items identified by AI platforms were thoroughly

[12] Wohlin (2014); Greenhalgh & Peacock (2005)

checked against Google Scholar to verify weblinks and citations, prior to reviewing the specific content of the sources.

Study selection and interpretation were guided by the authors' expertise, professional knowledge and experience in the main areas covered by this review (operational policing, police leadership and governance, community safety and well-being, emerging technologies, Canadian law and legislation, digital technologies, and the intersections of these domains from the standpoint of human and organizational considerations). Using social scientific and practitioner judgement, they iteratively examined and assessed the operational relevance and credibility of the available evidence.[13]

Key Informant Interviews

A set of exploratory, semi-structured, conversational interviews was conducted with a purposive sample of seven subject matter experts over video conferencing between February and May 2026.[14] These individuals had deep, overlapping, professional knowledge and experience in the following areas: police leadership and operational policing; technology-enabled law enforcement and investigations; technology-enabled crime; law and policy as related to policing and police reform; human rights and civil liberties; and transnational crime.

Prior to the interview, each participant was provided with background information on the project and the focus of the conversation. They were also assured that the interview would be confidential and that no comments would be attributed to any individual. They were given an opportunity to ask questions and asked for consent to audio-record the interview. All participants provided permission for the interviews to proceed as described. Each interview lasted approximately 45 to 60 minutes and was conducted by one or more members of the research team. Interview recordings were digitally transcribed.

Individual interview transcripts were reviewed manually to generate a provisional list of key themes. This was continued with thematic analysis at the group level to identify key insights and implications related to the focus of the project. Multiple passes were made through the interview transcripts to re-check and confirm saturation of the identified themes.[15] Insights from the literature review and the interviews were considered together in discussing the implications of AI for Canadian policing.

Expert Panel Review

A draft of this report was reviewed by a multidisciplinary expert panel in September 2026. It assessed the draft’s findings and recommendations for practicality, risk sensitivity, and scalability across services of different sizes. Feedback emphasized phased implementation, procurement safeguards, evidentiary disclosure, scope control, meaningful community authority, and support for smaller services. These themes informed the revisions reflected in this version.

---

[13] Moser & Korstjens (2017); Pawson, et al. (2003)

[14] Bamberger, et al. (2006)

[15] Charmaz (2000)

AI Use in Policing: Impacts and Implications

The following four areas capture the most consequential ways AI is affecting Canadian policing today and in the near term. Each presents both genuine opportunity and significant risk. Furthermore, each carries governance implications that police leaders and boards cannot treat as purely technical or operational matters.

1. Enhancing Effectiveness and Efficiency in Policing

AI's most immediate value to policing lies in its capacity to manage scale, complexity, and speed; the types of pressures that are increasingly defining contemporary police work. For police agencies, **the use** of AI can be thought of as an enabling infrastructure.

Organizational and Operational Efficiency

AI systems can automate routine and time-intensive tasks, including: interview transcription; report drafting; call triage (emergency and non-emergency); administrative workflows; and member training. Other applications include support for employee wellness, absence management and other HR-related functions.

A variety of AI-based systems are also in use to enhance cross-domain operational capabilities. As part of imaging systems, they can be used to enhance the utility of uncrewed aerial vehicles in search or surveillance operations or as part of the forensic reconstruction of serious motor vehicle accidents. In public order contexts, such as large demonstrations, sporting events, or major public gatherings, image recognition systems may be used to assist with crowd monitoring, identifying emerging safety risks, or locating individuals of concern. In joint national security operations, image recognition may be applied to critical infrastructure protection and counterterrorism activities.

These applications are advertised as being able to reduce staffing pressures, such as the cognitive load on officers and time on-tasks for a variety of personnel, to allow finite human capacity to be redirected toward judgement-intensive and community-focused work, without altering the core discretionary role of sworn members. These claims merit scrutiny. Administrative uses such as transcription or scheduling generally carry lower risk than operational uses that directly affect members of the public, and the two should be assessed against different thresholds.

Investigative Capacity

Complex crime investigations and prevention activities relying on evidence sorting and the fusion of intelligence data also lend themselves to a variety of emerging technologies. AI can make linkages that humans may not or could not have seen or made.

Advanced analytics, pattern recognition and image matching, and language-processing tools can accelerate investigations involving large datasets and digital evidence, such as financial records, video records, or open-source intelligence.

Internal police AI systems can ingest and analyze video data sourced from privately owned networked doorbell cameras, enabling algorithmic identification, tracking, and correlation of

individuals, vehicles, and events across distributed residential surveillance feeds spanning multiple locations and timeframes.

These technologies are being used to assist with the large volumes of data that are often involved in cases focusing on financial crimes, organized crime, cybercrime and child sexual exploitation. In the latter case, automated technologies may have the added benefit of helping to alleviate the potential for stress among human operators by lessening their exposure to traumatic imagery. AI platforms may also help to reveal patterns across several cases, supporting case clearance or cold case resolution.

Canadian courts are also signalling a growing interest in ensuring the responsible and transparent use of AI in a range of proceedings. Examples include: requirements to certify the authenticity of every factum or expert report submitted to court; requirements that parties disclose the use of AI in the preparation of court materials; and recent decisions involving the sanctioning of the irresponsible use of AI in courts, such as the use of faulty citations created by AI.[16] In *Ko v. Li* (2025), counsel filed a factum citing non-existent, AI-hallucinated cases. The court held that "counsel may not mis-state or misrepresent the law to the court whether by way of AI hallucinations or by any other means," and that counsel must verify the authorities they cite (para. 22). Hence, there is an expectation that professional organizations using AI tools will do so in informed and appropriate ways.

Resource Deployment and Risk Awareness

Data-informed forecasting tools may assist with strategic planning, staffing, situational awareness, risk management, and deployment decisions. Scenario modelling and forecasting tools may assist senior leaders with deployment planning and organizational risk management. Their value is enhanced through proper and disciplined use and an explicit recognition of potential sources of error or uncertainty.

A range of robotic technologies may be deployed during critical incidents to reduce risk to officers and members of the public while maintaining operational control. These uses are not aimed at increasing force but, rather, at reducing risk. Robotic tools extend officers' eyes, ears, and reach into dangerous spaces, helping teams gather information, communicate, and stabilize situations without immediate physical exposure.

Scenario | Implications

Scenario: An agency using image matching and categorization, within an investigative case management sequence, may improve and accelerate investigator reviews of evidence in child sexual exploitation cases, while limiting employee exposure to traumatic imagery. The determination of what constitutes appropriate evidence to put forward in support of charges remains at the level of the individual human investigator.

---

[16] Lockwood et al. (2025)

Implications: The use of AI can improve police effectiveness where it augments professional capacity, not where it displaces responsibility or obscures accountability. It adds value where it amplifies professional judgement; it creates risk where it substitutes for it.

---

2. Broader Regulatory Context

Police services deploying AI-enabled tools must navigate an interconnected legal and policy framework that spans: constitutional law; privacy legislation; criminal procedure; administrative law; and emerging digital governance regimes. The regulatory context is therefore multi-layered and evolving.

At present, AI in Canada is governed largely through privacy law; for private-sector entities, this is the *Personal Information Protection and Electronic Documents Act* (PIPEDA). Canada may soon legislate AI from the perspective of AI in general or AI via particular sectors (AI in healthcare, AI in education, etc.). There may be amendments to the Criminal Code to fill gaps that have been created by AI.

Constitutional Law

AI systems that influence surveillance, search, detention, risk assessment, or investigative prioritization may engage several sections of the *Canadian Charter of Rights and Freedoms*. These include sections: 7 (life, liberty, and security of the person); 8 (protection against unreasonable search and seizure); 9 (arbitrary detention); and 15 (equality rights).

Courts are increasingly assessing not only the outcome of police actions, but also the processes and technologies that shape decision-making. Algorithmic tools that affect investigative thresholds or evidentiary pathways must therefore withstand constitutional scrutiny for reasonableness, proportionality, and procedural fairness.

Privacy Law

Federal policing bodies are governed by the *Privacy Act*, while private-sector technology vendors are subject to PIPEDA. Municipal and provincial police services are subject to provincial privacy legislation and police services acts, with oversight by provincial information and privacy commissioners.

The Office of the Privacy Commissioner of Canada (OPC) has issued guidance emphasizing transparency, necessity, proportionality, data minimization and algorithmic accountability in the use of AI and facial recognition technologies. The Clearview AI controversy shows that relying on vendor assurances does not relieve police of their own statutory obligations.

In June 2026 the federal government introduced Bill C-36, the *Protecting Privacy and Consumer Data Act* (PPCDA). If enacted, this legislation will repeal Part 1 of PIPEDA to establish a modern federal private-sector privacy system for Canada. Bill C-36 aims to strengthen consent-based collection of personal information, and would create a new federal regulator, the **Digital Safety and Data Protection Commission of Canada** (replacing the OPC) to enforce online safety laws, oversee social media platforms and AI chatbots, and protect children from harmful

digital content. The Commission would have expanded enforcement powers, including the ability to impose significant administrative monetary penalties (AMPs).[17]

Criminal Procedure

Criminal procedure and evidentiary law are also implicated. AI-generated outputs (and the prompts used to generate the outputs) used in investigations may raise mandatory disclosure obligations under R. v. Stinchcombe, reliability considerations under expert evidence standards (e.g., R. v. Mohan), and challenges related to proprietary systems where defence access to underlying methodology is constrained. Courts have signalled that technological opacity cannot shield the Crown from its disclosure and fairness obligations. The Crown's argument that the technology should not be disclosed due to it being a protected investigative technique, is likely to be harder to sustain given the volume of publicly available documentation about policing technologies. This places a premium on contract terms. Where vendors will not disclose system logic, training data provenance, or validation results, or will not support expert testimony when evidence is challenged, the risk of a failed prosecution shifts directly to the police service and the Crown.

Administrative Law

Beyond traditional policing statutes, Canada's broader digital governance regime is relevant. The *Directive on Automated Decision-Making* (Treasury Board of Canada) applies directly to federal institutions and signals emerging norms around algorithmic impact assessments, explainability, transparency and human oversight.

Although municipal police services are not formally bound by the Directive, it provides a benchmark for responsible AI governance in public-sector contexts. Proposed reforms under the *Artificial Intelligence and Data Act* (AIDA), part of Bill C-27, introduced by the previous Trudeau government, indicated a future direction toward risk-based regulation of high-impact AI systems, including transparency and harm mitigation requirements. As an election was called prior to Bill C-27 receiving Royal Assent, the Bill did not become law.

Digital Governance Regimes

National security and cybersecurity legislation, including the *National Security and Intelligence Review Agency Act,* the *Security of Information Act*, and the proposed *Critical Cyber Systems Protection Act* (Bill C-26), which also died upon the 2025 election, highlight the expectation that federally regulated public institutions manage digital infrastructure risks, supply chain exposure, and cross-border data sovereignty issues with heightened diligence. AI systems integrated into policing environments expand the cyber-attack surface and may introduce jurisdictional exposure if hosted or maintained by foreign-domiciled vendors.

---

[17] Parliament of Canada (2026: https://www.parl.ca/legisinfo/en/bill/45-1/c-36; Osler, Hoskin & Harcourt (2026)

Scenario | Implications

Scenario: A mid-sized Canadian police service considers adopting an AI-enabled investigative analytics platform to improve efficiency and case prioritization. What appears to be a routine procurement decision quickly expands. The service must assess privacy legislation compliance, Charter implications, criminal disclosure obligations, records retention, cybersecurity risks, and bias embedded in historical data. Accountability, auditability, and community trust also arise, requiring board-level oversight and cross-functional review.

The service ultimately establishes a review process involving legal counsel, privacy officers, IT security, frontline supervisors, and community advisory representatives - recognizing that AI adoption is not simply about buying software but about embedding a socio-technical system into a public institution with constitutional and democratic obligations.

Key Implications: For police leaders, AI adoption is more than a procurement decision. It carries constitutional, privacy, evidentiary, and cybersecurity governance dimensions. Responsible acquisition requires early legal review, risk assessment, vendor due diligence, auditability, transparency mechanisms, and alignment with emerging federal AI norms. Legitimacy will depend not only on operational effectiveness, but on demonstrable compliance with Canada's rights-based regulatory environment and ongoing community engagement. Where IT and procurement functions sit within shared municipal services, that review must be anchored in formal agreements that leave the police service in control of its legal and evidentiary obligations.

---

3. Legal, Ethical and Practical Implications

Some of these technologies, notably applications involving employee or community surveillance, or behavioural prediction, are not without controversy. The legal and ethical challenges associated with AI in policing are not peripheral concerns; they are central to lawful authority and democratic legitimacy. The value of such tools depends on how clearly they are positioned as aids to human judgement rather than substitutes for it.

Technology alone doesn't create safety. Clear rules, defined thresholds, strong training, and visible oversight can help ensure that these tools strengthen public trust rather than weaken it.

Also important are the core assumptions and biases embedded within the design of algorithms, the data used to train and evaluate them, security, privacy and human rights safeguards. For example, the use of image recognition in operations or investigations occurs within a context of legal and policy constraints governing privacy, data retention, intelligence sharing, and oversight. Reports on the results of body-worn camera pilot projects are identifying their value in support of accountability, and in documenting the impacts of training and organizational culture on community relations and the interactions between both officers on-the-scene and members

of the public.[18] However, legislation, regulations and internal policies and procedures may struggle to keep pace with the introduction or adaptation of new technologies.[19]

Because these technologies can affect fundamental rights, such as freedom of assembly, freedom from unlawful search and seizure, mobility, and expression, their deployment in public order and national security environments generally requires careful governance, clear operational thresholds, auditability, and independent review to ensure proportionality and legitimacy. All of these considerations should be part of the lifecycle of developing business requirements, selecting vendors, acquiring tools, training, supervision, monitoring, governance and application retirement. These elements impact not only the assessment of the overall validity and functional utility of a given solution, but also the agency's legal liability and its social licence to use the tool in question.

Recent polling by the Angus Reid Institute identified that the Canadian public holds significant skepticism towards AI technologies, while one-and-a-half times more workers believe the effect of AI will be profoundly negative than those who believe it will be more positive than not.[20]

This skepticism extends to the question of who can be trusted to govern A.I. A separate Angus Reid survey found that 68% of Canadians believe government needs to heavily regulate it even if it slows development. Fully 74% of Canadians doubt that any government can keep pace with AI technology. For police leaders this is a relevant baseline. The public is not simply wary of AI tools themselves but doubtful that any institution, including government, can adequately provide oversight.

Bias, Discrimination and Systemic Risk

AI systems trained on historical data risk reproducing past patterns of over-policing, unequal protection, and systemic bias. This is particularly acute where predictive tools and biometric technologies are concerned. For example, tools trained and tested on data sets involving limited human diversity or biased historical data may result in "false positive" matches or invalid predictions of risk.

A well-known and often-repeated example was the application of the Clearview AI system across several Canadian police services. This became a focal point for broader concerns about police use of facial recognition technologies, particularly among Black, Indigenous, and other marginalized communities. Community advocates and legal experts expressed concern that deploying a tool built from scraped online images, without individuals' knowledge or consent, could deepen distrust in policing, especially in communities that already experience disproportionate surveillance.[21] Without rigorous auditing and contextual safeguards, AI may formalize inequities under the appearance of technical neutrality. In February 2021, the OPC released an investigative report into Clearview AI's system. The report identified significant

[18] e.g., Bachar (2020) re Durham Regional Police Services; Peel Regional Police, personal communications

[19] Recently, Burger King received negative publicity related to the implementation of AI headsets for customer service personnel that are set to monitor their interactions with customers (Kaye, 2026)

[20] Angus Reid Institute (20 November 2025; 08 June 2026)

[21] Battagello (2020)

shortcomings, ranging from accuracy and data collection to a real risk of harm.[22] A further concern was that police had submitted images of minors. Many jurisdictions have strict rules, codified in law, around the identification of minors involved with the criminal justice system. By inputting images of minors, a permanent record may have been created with no right to erase the image from Clearview's database or prevent it from being used for training.

Unrevealed biases within AI systems may also lead to the over-policing of racialized and marginalized communities. This constitutes a unique technological barrier to an informed community, able to engage their police service in evidence-based discussions for improved performance and accountability. It also runs contrary to recent work by Canadian police agencies to strengthen community relations and public trust through efforts such as race-based data programs and human rights-based policing initiatives.

This highlights both real and perceived risks: even where technology is used for legitimate investigative purposes, its association with mass data harvesting and unclear practices can undermine confidence in police judgement and oversight.

From a governance standpoint, the Clearview example illustrates that the legitimacy challenge is not solely about technical accuracy or investigative value. It is also about transparency, legal authority, proportionality, and public understanding. Community reaction to AI tools may hinge less on operational effectiveness and more on whether agencies can clearly articulate the legal basis for use, demonstrate safeguards against bias, and show that oversight mechanisms are robust. In some cases, it will also be important to address potential misalignments with national commitments to reconciliation and Indigenous rights. It also means giving affected communities genuine input and influence over certain kinds of technology adoption decisions. Fear of surveillance is itself a harm - it can deter people from reporting crime, seeking help, or exercising lawful rights, even where no misuse occurs.

For example, the use of AI-based technologies by police may, in some cases, conflict with parts of the Truth and Reconciliation Commission of Canada's (TRC) Calls to Action[23] and/or the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP).[24] This could include: concerns about mass surveillance; the need for special attention to the rights and needs of Indigenous peoples (such as the use of appropriate training and evaluation data sets); the right to access fair, just and accountable procedures, or the adequacy of systems to protect Indigenous rights, as well as the impact on traditional territories of resource extraction and climate disruptions tied to the development and operation of AI based systems (e.g., data centres).[25]

---

[22] Office of the Privacy Commissioner of Canada (2021)

[23] Truth and Reconciliation Commission of Canada (2015)

[24] Office of the United Nations High Commissioner for Human Rights and Asia Pacific Forum of National Human Rights Institutions (2013)

[25] In 2025, following significant public outcry, a Hamilton, Ontario planning tribunal rejected a proposal to use a plot of land on which to build a large-scale AI data centre.

The Clearview controversy, and potential tensions involving national commitments to TRC and UNDRIP, serve as reminders that AI deployment decisions are as much about factors related to community trust and social licence as they are about enhancing investigative capabilities.

Many discussions of the use of AI in consequential applications consider it important that fully autonomous systems not be deployed. When humans are "off-the-loop", the various biases and errors these systems are prone to produce would go undetected.

One school of thought for dealing with this involves having a human involved in reviewing and approving AI-generated recommendations before action is taken. This is known as a "human-in-the-loop" (HITL) situation.[26] Another approach, known as "human-on-the-loop" (HOTL), involves systems that work with more autonomy, but that are subject to human supervision, with monitoring, intervention and override opportunities.[27] These approaches are intended to reduce the risks associated with autonomous systems and reinforce human accountability.

Some commentators believe that fast-tempo, information-dense, operational situations, such as military combat, will increasingly shift from HITL to HOTL.[28] However, effective oversight requires more than assigning a person to a process. Human reviewers need the authority, training, information, supervision, support and time to evaluate AI outputs critically and challenge recommendations when appropriate. Without those conditions, oversight can become a box-checking exercise rather than a meaningful safeguard.

For police services, the key question is not only where AI can add value, but also where human judgement should remain central to operational, investigative, and administrative decisions, contexts that differ significantly from military combat. But human oversight and governance are necessary but not sufficient on their own. The presence of a human in the loop does not guarantee meaningful control.

Facial recognition illustrates the risk well. As one interviewee put it, "the [individual] cognitive bias that might accompany a determination that you have a facial recognition match is pretty high." AI-generated outputs can anchor, rather than merely inform, the judgement of a reviewer working under time pressures, particularly where the tool presents a confident-looking match or summary. Mitigating this requires pairing human review with explicit training on automation bias, a documented rationale whenever a reviewer accepts or overrides a consequential AI output, and supervisory audits that sample reviewed decisions to confirm that primary sources are still being examined.

Data Confidentiality, Integrity, Availability and Output Quality

When AI platforms are built on top of other AI platforms, code or algorithms that themselves contain errors, misrepresentations or cybersecurity vulnerabilities, the resulting tools will be more susceptible to producing faulty output, which is sometimes referred to as "AI slop",[29] as

[26] Crootof, et al. (2023)
[27] Ibid
[28] Milmo & Down (2026)
[29] Bender & Hanna (2025)

well as to malicious penetration. Where application output is used within highly consequential processes, the results for systems, organizations, neighbourhoods and individuals may be serious.

A well-known example of the dangers of AI slop within consequential processes was the use of a chatbot by the US National Eating Disorders Association to replace its hotline operators[30]. The results included feedback and advice to callers that was not only harmful or insulting but, in some cases, dangerous. Using synthetic media platforms to replace human operators in emergency/non-emergency dispatch centres may entail similar risks for faulty outputs. While AI-enabled call triage may be deemed necessary in order to better enable the deployment of finite resources to handle the volume of lower-priority calls, the potentially serious implications of misunderstood and misclassified calls should not be ignored.

Crowdwork

AI tools are often presented as being strictly technological solutions related to human problems of efficiency, productivity, or cognitive load (analytic or decision-making burden). As Bender and Hanna (2025) observed, not every AI tool that claims to be "autonomous" is fully free from human input or supervision.

In some cases, these technological tools are based on an infrastructure that requires offshore labour to operate. For example, one rideshares company operates driverless vehicles that are often thought of as completely autonomous. However, rider assistance calls fitting certain parameters are shunted to human operators, some of whom are contractors working in the Philippines[31]. A 2026 investigation by Swedish newspapers suggested that Meta was using video from Ray Ban Meta AI glasses, its popular wearable glasses, and sending it to Kenya for human reviewers to ensure the AI algorithm was properly trained.[32] It was alleged that reviewers were exposed to videos that caught sensitive and intimate moments of users within their lives.

These situations have three practical implications for end-users. First, both the technology and the human operators who support it may be vulnerable to errors. Second, while unlikely, there is a possibility that these workers could be compromised or coerced, potentially creating security risks for the organization. Third, community stakeholders may question the use of tools that rely on low-wage offshore labour - particularly if those workers have limited ability to raise concerns about system failures or ethical issues.

Human Rights: Privacy, Civil Liberties and Charter Compliance

AI-enabled surveillance, data aggregation, operational scope creep and behavioural inference call for attention to privacy concerns. The scale and persistence of data collection enabled by AI can alter the balance between state power and individual rights, even when individual tools are lawfully deployed. AI intensifies an important tension around the need to support victims and

[30] Noguchi (2023)
[31] Sims (2026)
[32] TechCabal (2026)

pursue justice while safeguarding Charter rights, within the boundaries of the law and democratic accountability.

Multiple interviewees and other expert views captured in this review highlighted risks of "slippery expansion" once AI capabilities are acquired. Examples included:

- Digital evidence (forensic examination to metadata) analytics repurposed for live surveillance;
- Call screening tools potentially expanding to emergency calls;
- Historical controversies and emerging concerns (e.g., Clearview AI, IMSI catchers,[33] SMS Blaster,[34] video doorbells).

Their concern was less about initial deployment and more about the absence of structural constraints preventing incremental expansion. To remedy this, every deployment should carry a defined, documented set of objectives and accompanying description of the intended scope of application, and any material change, such as extending a tool from recorded to live video, or from non-emergency to emergency calls, should trigger a new risk assessment and fresh approval.

AI-enabled surveillance and data aggregation may challenge existing privacy norms and legal thresholds. For example, privacy protections enabled by the circumscribed parameters of specific investigations may be eroded as a result of aggregating data across investigations. Similarly, if the use of image recognition technologies approved for use with static images is extended to live video feeds, this may violate not only some agencies' internal guidelines and governance parameters but also undermine public confidence. At a technical level, untested applications may also introduce evidentiary or Charter challenges.

A further issue concerns the identity and reputation of a technology vendor. In some cases, a platform may have a broad array of applications beyond law enforcement, including military intelligence, enemy targeting, autonomous weaponry and other functions across a range of clients and operational environments. To the extent that these applications or clients are a focus of concern for community members, their adoption by a local police service may cause alarm among certain groups in the community. The prevalence of certain platforms and the benefits of interoperability across agencies can make these decisions difficult ones for police leaders and their governing bodies.

Chiefs must consider not only whether individual tools are lawful, but whether cumulative use or vendor reputation alters the balance between state power and individual rights in ways that courts or oversight bodies may later reject, or that may undermine public trust.

[33] International Mobile Subscriber Identity (IMSI) technologies are tools that identify, locate, and may interact with mobile phones by using the phone's connections to nearby cellular networks.

[34] Fanfair (2026)

From the standpoint of the Charter, AI adoption, especially when it impacts on privacy, should be able to be explained in terms of demonstrable public benefit. Efficiency alone is unlikely to constitute sufficient justification from a Charter perspective.

The report from a recent roundtable on AI and policing, convened by the University of British Columbia, positioned necessity, proportionality and demonstrable public benefit as foundational for effective governance of AI tools within the policing context.[35] In this sense, demonstrable public benefit means not just the fact of improved public safety, but the capacity to explain to members of the community how a particular technology works to advance their safety and how their privacy and liberty are protected in the context of its use.

Unproven Performance Claims

Vendor claims about the performance of their technologies may be difficult for lay persons to understand and verify. There is a need for clear, reliable and explainable methods to assess the accuracy of the claims and to appropriately examine the algorithm. The expertise required of procurement personnel and those in organizational leadership and governance, who are ultimately accountable for the choice to deploy an application, increasingly needs to span technical knowledge, legal knowledge and human rights, community knowledge and practice domains. Even when evaluation metrics are published by a vendor, it is not always clear what the bases for measurement were, nor how these parameters relate to the ways in which the system is intended to be used.[36]

Participants of the UBC roundtable on AI and policing observed that independent empirical validation of many police AI technologies remains limited, particularly with respect to reliability, accuracy, effectiveness and demonstrated necessity.[37] They suggested that, while efficiency gains are frequently highlighted, operational results have been inconsistent in practice. In some cases, AI tools were found to introduce additional human resource burdens, such as reviewing and correcting inaccurate report content. At the same time, they referred to accumulating evidence that the accuracy of facial recognition systems is improving, especially under controlled conditions.

Besides generating inaccurate content, AI systems pose the risk that employees may stop reading original source files altogether. One participant described a situation where the pressure of a high volume of pending case reviews led decision-makers to rely on AI-generated summaries as the basis for their determinations. These types of behavioural shifts can be just as problematic as function creep in introducing errors, and sidelining the use of human judgement, in critical areas of organizational performance. Supervisory audits that test whether primary sources are still being reviewed are an essential counterweight.

Effective performance also entails reliable security features that span the design of systems, the structure of system functions (including data residency and how and where data-in-use transits, as well as alignment to cybersecurity standards and relevant legislation) and organizational

[35] Perrin (2026)
[36] Bender & Hanna (2025)
[37] Perrin (2026)

policy and process safeguards related to system and data confidentiality, integrity and availability. In situations involving one or more subcontractors, the vulnerability of a vendor's product may have as much to do with the vulnerability of upstream suppliers as with the vendor, itself. Finally, the sovereignty of systems serving Canadians has also begun to be an important consideration across multiple sectors. This touches on issues related to concerns about both availability and resistance to misuse.

Algorithmic Transparency and Explainability

Many AI systems operate opaquely, making it difficult for affected individuals or police organizations themselves to understand how outputs are generated or how automated systems are controlled. This creates challenges for disclosure, judicial scrutiny, and meaningful oversight. Tools that cannot be explained to courts, boards, or the public create legal and reputational exposure. In these situations, explainability is not a technical preference; it is an operational requirement.

AI is often described as a "black box," but the real governance challenge for policing is not just whether a vendor is willing to open the box, but whether the agency is in a position to understand what is disclosed.[38] Two important concepts are that:

- transparency is not a single act of disclosure, it is a layered architecture; and
- access to information is not the same as explanation (Wischmeyer, 2019).

This begins with public acknowledgement that AI systems exist and are shaping operational decisions. It extends to meaningful explanations for individuals whose rights or liberty may be affected. It culminates in something more consequential: empowered oversight bodies, courts, and internal governance structures with the technical literacy and authority to scrutinize models, interrogate training data, and test system performance. In other words, explainability is not merely a feature of the software, it must be an embedded and developed capability within the organization's culture and practices.

For police services operating within a constitutional liberal democracy, this is important. Transparency cannot stop at vendor assurances, model summaries, or simplified counterfactual explanations.[39] Effective governance requires that the organization (and its oversight bodies) be able to examine how the system was built, what data it was trained on, how accurate it is, and whether it produces fair and lawful outcomes. A simple explanation helps people understand a decision. Full institutional transparency ensures the decision deserves to be trusted.

Where AI systems are used to guide resources, such as patrol deployment, investigative prioritization, risk assessments, or identification processes, the stakes are not the same as in

[38] "Box opening" might occur at the point of vendor selection, or on an as-required basis, in the context of court proceedings.

[39] e.g., "If the risk score had been lower, the case would not have been flagged." This kind of counterfactual explanation seems clear and concrete, but it does not show the full picture. Modern AI systems rely on many interacting factors, historical data patterns, and statistical weights. Highlighting one variable obscures deeper issues such as biased training data, questionable thresholds, systemic patterns built into the model, or restricted evaluation data sets. A simple explanation is not the same thing as accountability.

everyday commercial applications. Instead, they have legal implications. If secrecy claims, proprietary barriers, or institutional capacity gaps prevent meaningful review, the issue is no longer about technology; it is about legitimacy. The question for police leaders is not whether AI can enhance operational effectiveness - it is whether their governance architecture, from procurement, oversight, documentation, judicial accessibility, to public communication, is robust enough to ensure that algorithm-based uses remain accountable to the rule of law and deserving of public trust. When information sharing (whether about AI-based systems or about individuals whose information is held in these systems) involves national boundaries,[40] different laws and conventions may apply. This further adds to the complexity of understanding and managing these processes.

Transparency must also be calibrated. Public disclosure at the level of the tool (what it is, what it is used for, and how its outputs can be challenged) is essential to legitimacy. Method-level detail that could compromise investigations can instead be disclosed to an independent oversight body with the clearance and expertise to examine it. Both layers are necessary; neither is sufficient on its own.

Labour Relations and Occupational Health and Safety

AI introduces a new set of workforce considerations alongside its operational benefits. Decisions about how AI is deployed, what data it uses, and how its outputs are applied can have significant implications for employee trust, labour relations, morale, and organizational legitimacy.

In 2026, London's Metropolitan Police Service (UK) began piloting AI tools supplied by Palantir to help identify officers who may be at risk of misconduct.[41] The system analyzes internal workforce data, such as attendance, public complaints, use-of-force reports, and other indicators of the potential need for administrative intervention, so that these concerns can be addressed earlier. While supporters described the approach as a way to strengthen accountability and rebuild public trust following a series of misconduct scandals, concerns have been raised about transparency, privacy, fairness and the potential for inaccurate or biased results; the Metropolitan Police Federation criticized the approach as "automated suspicion".[42]

This example underscores the need for clear governance, oversight, validation and advance employee and union engagement when using AI to support internal professional standards and workforce integrity functions. If AI-generated performance metrics were to be used, flagged employees would be expected to question the basis for being identified. In the event of termination and subsequent litigation, experts will debate the merits of the system underlying the decision to terminate. This would risk exposure to both potential liability and to reputational damage. Police services considering the use of AI would benefit from engaging early with labour representatives and employees to address concerns about workplace surveillance, job impacts,

[40] For example, this could involve where and how data is transferred or stored; or foreign ownership of a particular platform or tool deployed Canada.

[41] MPS Confirms the Use of AI (2026)

[42] Booth (2026)

decision-making authority, and the appropriate use of AI in performance and operational environments.

Similarly, a 2025 study of how AI-assisted body-worn cameras related to officers' perceptions of procedural fairness found that officers were more accepting of AI oversight when they perceived the process as fair and transparent.[43] The findings suggested that how AI tools are introduced, not just what they do, is central to whether police workplaces adopt them without conflict.

Other research looking at the introduction of AI-driven body cameras in two Arizona police departments found that automated positive feedback improved officer acceptance.[44] This contrasted with experiences in Seattle, Washington, and Vallejo, California where, for various reasons, police unions opposed the adoption of AI body camera analytics and where at least one department let its contract expire in the wake of union pressure[45].

Police officers may also become targets of multimodal information campaigns. This may involve the use of visual and/or auditory records to call to account or discredit an individual either as a result of conduct that was honestly understood by a bystander as inappropriate, or in the context of broader efforts to maliciously target police for other reasons. Officers have already faced doxxing and coordinated online abuse (including technology facilitated gender-based violence), with identified effects on their safety, well-being, and willingness to take on public-order roles.[46]

Synthetic media, such as the use of deepfakes depicting officers involved in misconduct, is likely to intensify these risks. The result will be pressure on organizations to find new ways to support the occupational health and safety of their members.

Scenario | Implications

Scenario: A police service considering an AI system trained on historical policing data determines that false positives represent at least a moderate risk. Media reports from other jurisdictions indicate the system encodes past enforcement patterns, including over-representation of certain communities, prompting local concern. Before proceeding, the agency requires the vendor to provide a full explanation of the system's design, evaluation, key features, and contextual safeguards - ensuring potential sources of bias are identified and corrected rather than institutionalized. That information is reviewed collaboratively with community representatives, the police board, and external technical and legal expertise.

Implications: AI that cannot be explained, challenged, or governed consistent with Charter values will ultimately undermine both police authority and effectiveness. Multi-layered transparency is essential to effective governance, management and operational capability. Where that capacity does not exist in-house, external expertise should be acquired, and

---

[43] Adams (2025)
[44] Watts, et al. (2025)
[45] Du Sault (2023)
[46] Huey et al. (2025); Watson (2023)

meaningful mechanisms for community engagement must be built into the process. The agency should also be prepared to walk away and seek alternatives if those conditions cannot be met.

---

4. Threats, Risk and Security

AI introduces new categories of operational and institutional risk that extend beyond traditional technology management. It increases institutional exposure in three main ways.

Cyber Vulnerabilities and Adversarial Exploitation

AI systems concentrate sensitive data and create attractive targets for cyber intrusion, manipulation, and theft. They may also be the subject of non-human interference. Concentrated datasets and models become high-value targets. Cyberphysical systems can be targeted, deceived, or reverse-engineered by sophisticated adversaries. This includes spoofing biometric systems, poisoning datasets, or exploiting automated decision thresholds. Criminal and hostile actors may exploit (e.g., manipulate data to compromise an investigation - also known as "data poisoning"), deceive, or mirror police AI tools.

A 2025 report by the United Nations Office on Drugs and Crime (UNODC) described how criminal groups are combining generative AI, automation, deepfakes, and cybercrime infrastructure to create highly scalable fraud operations.[47] Similar to what has been seen in the field of cybersecurity, where "ransomware-as-a-service" has been available for some time, we are now seeing "crime-as-a-service" models in which specialized actors provide AI-enabled fraud tools, synthetic identities, malware, money laundering services, and other capabilities to the criminal marketplace. This trend suggests that sophisticated AI-enabled offending may become accessible to a much wider range of offenders over time.

Another emerging threat area that merits attention from police leaders is the use of generative AI to facilitate child sexual exploitation, sextortion, and the production of synthetic child sexual abuse material (CSAM). AI is being used to make hyper-realistic deepfake sexual imagery for those who consume CSAM, as the basis for sexploitation, or to automate grooming tactics that were previously labour-intensive. Offenders can create synthetic abuse content by manipulating images scraped from social media or other online sources, including images of real children with no prior experiences of victimization.[48]

Experts warn that these technologies may lower barriers to offending, increase the scale of exploitation, complicate victim identification, and create new investigative challenges as synthetic and authentic imagery become increasingly difficult to distinguish.[49] Currently, the federal government has indicated that laws specific to combatting AI-generated deepfakes are being considered. In the EU, Denmark is exploring whether copyright law amendments could

---

[47] UNODC (2025)
[48] United Nations Children's Fund (UNICEF) (2025)
[49] Internet Watch Foundation (2026)

solve the issue by providing an automatic copyright to your likeness from birth.[50] From a policing perspective, the issue highlights how AI can amplify existing forms of exploitation while creating entirely new victimization pathways.

Looking ahead, the most significant AI-related threat to policing may not be any single technology, but the growing accessibility of AI-enabled criminal capabilities. Deepfakes, automated fraud, synthetic identities, cybercrime tools, and AI-assisted social engineering are reducing the expertise and resources required to commit sophisticated offences. As these capabilities become embedded within organized crime networks and criminal service marketplaces, police services will need to anticipate faster-evolving threats, higher volumes of technology-enabled crime, and increasing challenges in verifying the authenticity of digital evidence.

Compromised data used to train or weight algorithms, or model tampering, can quietly distort outputs in ways that are difficult to detect. For example, the training data used by AI platforms to generate programming code are often based on repeats of common coding patterns which may be less secure.[51]

System breaches at the service provider, cloud or customer level may lead to the exfiltration of protected information, damage to the privacy or safety of individuals, or the malicious disruption of data holdings ("data poisoning"), potentially compromising investigations, prosecutions and public trust.[52]

The use by police personnel of unapproved technology, often referred to as "shadow IT", can be thought of as both an extension of formal and informal "Bring Your Own Device" (BYOD) practices that are now common across many organizations, and the tendency of action-oriented, problem-solvers to find practical workarounds to challenges they encounter on-the-job. The use of generative AI applications[53] has exploded across many workplaces in just the past few years[54].

As a result of their widespread presence as apps on personal devices, they could be the next wave of "BYO", what we might call, "BYOAI". The use of shadow IT has been identified as an enterprise risk touching on organizational security, data privacy and individual employee risks.[55] The UBC roundtable on AI in policing identified the experimental use of unapproved AI by police personnel as potentially introducing a variety of risks to police services and communities, including the potential violation of privacy rights. [56] The topics of AI-enabled security and the security of AI applications are very much developing spaces. They require ongoing input from

---

[50] The Associated Press (2025)

[51] Bender & Hanna (2025)

[52] These may include any number of cybersecurity or basic information security breaches, such as unauthorized or illicit/illegal access by insiders, or various cyber exploits carried out by external actors (e.g., social engineering, credential compromise, malware infiltration or exploitation of vulnerabilities).

[53] These include: OpenAI's Chat GPT; Anthropic's Claude; Google's Gemini; and Microsoft's Co-Pilot, among others.

[54] Hahmann & Lovei (2026)

[55] Deloitte (2025)

[56] Perrin (2026)

both technical experts and those with expertise on the human implications of improper or insecure use. Acceptable-use policies should therefore name unapproved AI tools explicitly as prohibited "shadow AI," and pair that prohibition with approved, secure alternatives so that the demand driving these workarounds is met.

A more recent threat has emerged related to the activities of misaligned AI agents. Misalignment occurs when the actions of one or more AI agents go beyond their design parameters leading to unexpected, unsafe, or unauthorized outcomes. For example, media reports from 2026 described how a group of Anthropic's agents acting independently of human direction sent a fake homicide tip to the Philadelphia Police Department.[57] This was accompanied by a series of media reports describing unauthorized actions of agents from Anthropic and OpenAI that had broken out of their testing environments and hacked into a number of networks, in order to harvest information that the agents' algorithms predicted would help them succeed in an evaluation task.[58] The agents did not just act alone; they aligned their activities in several work groups that were not orchestrated by humans.[59] Beyond simple misalignment are concerns that AI agents running on IoT devices may themselves, in some cases, become new attack vectors. New research suggests that self-replicating systems, powered by the energy of the devices they reside on (e.g., everyday appliances), could lead to significant damage across multiple domains.[60]

In the wake of this growing catalogue of incidents and potential threats, a growing chorus of international experts have called for a slower pace, if not a moratorium, on AI development until stronger safeguards and other conditions have been developed and put in place.[61]

Organizational Capture, Systemic Risk and Canadian Data Sovereignty

As AI tools become embedded in workflows, police agencies may become dependent on proprietary vendor-owned systems they do not control and cannot fully audit. This creates long-term strategic risk, particularly where vendors provide limited or no audit information or long-term control, and where they operate across jurisdictions with differing legal standards and expectations and, sometimes, opaque accountabilities. This is a form of organizational capture. Capture can also deepen after a contract is signed. For instance, without notice, vendor-driven updates can change how a model behaves; a change in vendor ownership can alter who has access to data and under which jurisdiction; and escalating licensing and integration costs can make exit impractical. Contracts should therefore require advance notice and re-validation before updates are deployed operationally, protections in the event of a change of ownership, cost controls, and enforceable exit and data-portability rights.

---

[57] Conger, et al. (2026)
[58] Bhuiyan (2026); Tuquero (2026)
[59] OpenAI (2026)
[60] Guan, et al. (2026)
[61] Castaldo & Posadzki (2026); Cecco (2026)

Where an application is subject to either cybersecurity risks or foreign ownership, participating police agencies may become part of a systemic vulnerability present across the Canadian law enforcement community.

China's *National Intelligence Law* (2017) requires organizations and citizens to assist and cooperate with state intelligence activities in accordance with the law,[62] a provision that has been cited in international discussions about data security, supply chain risk, and the governance of technology providers operating across jurisdictions. Canadian security and intelligence authorities have, in some documented cases, assessed Chinese technology platforms and equipment as posing potential security or supply-chain risks, particularly where data access, network integrity or foreign legal obligations are at issue.[63]

More recently, concern has grown that many enterprise platforms and digital tools used in Canada are US-owned, or route data through the United States, even when that data is nominally stored in Canada. This lack of sovereignty over critical technologies represents a potential set of operational and privacy risks to the Canadian law enforcement community, in the event of unilateral foreign policy or regulatory actions by the United States.

Organizational capture is also compounded by a capacity gap that runs across the sector. Many Canadian police agencies, particularly small and mid-sized ones, lack in-house technical expertise to evaluate AI systems on their own terms. As one interviewee put it, "police forces… are not engineers… they inevitably end up going by the vendors' claims on how their tools work… there's no mechanism for scrutinizing those in the ways that you would need to." Knowing what questions to ask of a vendor, and what counts as a reasonable answer, is not an intuitive skill for police personnel and for board and commission members alike.

In Ontario alone there are 45 police services of vastly different sizes and capabilities - even larger, better-resourced agencies can struggle with vendor capture, platform dependence, and independent technical evaluation. Without shared regional, provincial, or national evaluation mechanisms, agencies risk repeating one another's failures, incurring avoidable legal exposure, and damaging sector-wide legitimacy. Perrin (2026; Perrin, et al., 2026) identifies untapped potential in existing oversight bodies already present in most Canadian jurisdictions (e.g., Ontario's Inspectorate of Policing, police service boards, British Columbia's Director of Policing and Law Enforcement Services), describing this as a "multi-layered oversight ecosystem" that could be anchored into one or more provincial frameworks for AI governance rather than building from scratch.

In the interim, existing resources such as the IACP Connector program offer agencies a practical starting point. The Connector program connects police agencies with peers who have navigated similar procurement decisions and provides access to shared experience, profiles, vendor information, and AI-specific guidance.[64]

---

[62] Standing Committee of the National People's Congress of the People's Republic of China. (2017)

[63] Government of Canada, Ministry of Innovation, Science and Industry, & Ministry of Public Safety. (2022); Canadian Centre for Cyber Security. (2024)

[64] IACP (2026)

Capacity constraints are compounded where police IT and procurement are delivered through shared municipal services. In those environments, key decisions may be made outside the police service, yet the legal, privacy, and evidentiary consequences remain with the chief and board. Formal service level agreements and a cross-governance framework can clarify roles while keeping the police service in control of its compliance obligations. Over time, smaller services would also benefit from shared services, a vetted technology inventory, and access to legal, privacy, and academic expertise that no single agency of their size could sustain.

At the recent June 2026 G7 summit, senior executives from leading AI firms were reported to have warned that current governance arrangements are struggling to keep pace with advances in AI capability. The discussion highlighted the increasing concentration of AI power and expertise, as well as implications for national security, public oversight, and democratic accountability. Their remarks reflected a growing concern that technological development is outpacing existing governance, regulatory, and accountability mechanisms.

Platform Dependence, Decision Shaping and Practice Drift

Critical thinking, along with integrity and practiced technical skills, is a core requirement of all members of a police service. It enables split-second decisions by front-line officers who also must articulate the rationale for their use of powers and authorities. It also underlies the capacity of senior leaders and board members in their exercise of organizational oversight, responsibility for strategic outcomes, and discharge of fiduciary responsibilities.

An over-reliance on a particular technology or platform can shape how people think about their work. Over time, problems and opportunities may become framed in terms of what the system can measure or automate, narrowing how issues are defined and limiting more creative or critical solutions. A related problem involves the framing of social problems in ways that are used to justify the adoption by police of AI technologies.[65] A recent study of the use by Dutch police of technological solutions to address complex public safety issues found that certain potentially problematic aspects of these tools could be better managed through an ongoing process of dialogue with a diversity of stakeholders rather than thought of as a one-time decision that may be driven by pressure from vendors or internal stakeholders who are technology boosters.[66]

Police practices may also begin to evolve, following the introduction of powerful AI-based tools. A multi-year, field-based study with the Los Angeles Police Department, examined its use of various Palantir data platforms. It was found that, in some cases, the technologies accentuated existing police practice while, in others, the tools were associated with a shift of police surveillance practices towards those that more closely resembled military intelligence.[67] Personnel were seen to be drifting, in unintended ways, from information gathering based on probable cause towards activities anchored to predictive processes.

---

[65] Jacobs, et al. (2024)

[66] Ibid.

[67] Brayne (2017), reported in Crawford (2021)

In Canada, police do not need probable cause simply to conduct predictive intelligence. The legal trigger is whether the activity intrudes on a reasonable expectation of privacy under section 8 of the Charter. Analyzing lawfully obtained data, open-source information, or historical crime records is generally permitted. But when predictive work involves a search or seizure, such as accessing private digital data or conducting intrusive surveillance, police must have or obtain judicial authorization, typically based on reasonable grounds to believe an offence has been or will be committed. Although the core issue here is privacy intrusion, not whether an activity is labeled predictive, police leaders should anticipate and guard against unintended drift from lawful information gathering toward more legally questionable practices, particularly where powerful AI tools make deeper or more intrusive analysis easy, scalable, and routine.

Platform dependence can also stem from inadequate knowledge transfer, training and other capability-building processes that help ensure that the vendor's client is able to use the platform independently, without resorting to expensive consulting services. Similarly, it is well known among technology vendors that, once a client has become accustomed to a platform, it is very difficult for them to exit. Rising costs reinforce this effect - the longer an agency relies on an integrated platform, the more expensive it becomes to leave.

On a more positive note, platforms may enable new forms of collaboration across business and operational units, or between agencies. At the same time, they may structure how this happens in ways that constrain more organic and "out-of-the-box" approaches to collective problem solving, especially where this involves collaboration or partnership with community organizations or residents.

Scenario | Implications

Scenario: A Canadian police service adopts an integrated AI platform consolidating records management, analytics, digital evidence, video feeds, and case prioritization into a single dashboard. The gains are immediate: faster investigations, more data-informed deployment, reduced administrative friction, real-time command visibility. The system becomes the organization’s connective tissue.

But as reliance deepens, architecture begins to matter as much as outcomes. The platform's hybrid cloud environment and multiple vendor integrations expand the cyber-attack surface. Proprietary models and workflows make it difficult to benchmark alternatives or independently test assumptions. Vendor dependence grows as internal expertise erodes. Officers grow accustomed to automated risk scores and algorithmic flags; what gets prioritized increasingly reflects what the system was designed to measure.

Gradually, decision-shaping replaces decision-support. Patrol patterns mirror historical data embedded in predictive models. Investigative attention clusters around system-generated probabilities. Community stakeholders begin to question whether the platform reinforces legacy biases or shapes enforcement priorities in ways that are difficult to see or challenge.

Implications: AI platforms do not merely modernize existing tools. While often extending operational capability, they can rewire institutional behaviour and expand the organizational threat surfaces.

The issue for leaders is not whether AI improves operational performance, but whether governance structures, cybersecurity resilience, procurement safeguards, and internal expertise are strong enough to keep the organization in control of the technology, rather than the reverse. AI adoption is an enterprise-level decision carrying risk management and fiduciary responsibilities at the board and executive level. The strategic question is not adoption alone. It is also about maintaining autonomy, resilience, and public legitimacy. The governance of AI in policing cannot be a self-regulating process; it requires the ongoing, structured, and substantive engagement of the community.

---

5. Criminal, Ideological, and State Actors

AI amplifies online harms, such as disinformation, enabling them to reach scale far more rapidly. Canadian police services should anticipate:

- **AI-enabled fraud and impersonation**, including deepfake extortion and business email/system compromise.
- **Accelerated radicalization and disinformation**, targeting both the public and police personnel. Malicious information operations may seek to exploit existing social cleavages or anticipated labour dislocations caused by the replacement of workers by AI. According to the World Economic Forum's Future of Jobs Report 2025[68], AI and related technologies are expected to displace millions of jobs by 2030. While these disruptions will create new opportunities (based on skills like creativity, resilience, and problem-solving in the workforce), they may also dislocate many from the labour market, driving economic and social insecurity.

  Various ideological movements are likely to seek to capitalize on this to incite social unrest, in order to destabilize democratic processes and institutions. Increasingly, deepfakes are being developed and amplified across social media platforms by hostile actors and geopolitical adversaries as part of their online disinformation operations.[69]
- **Use of AI-enabled technologies to propagate abuse**, such as grooming prospective victims of child sexual exploitation and human trafficking, online sexual exploitation and abuse (e.g., technology-facilitated violence against women and child sexual abuse material using deepfake and nudification apps) and intimate partner violence.
- **Hybrid threats**, where AI-enabled cybercrime, disinformation, and national security concerns converge, including the use of "soft violence" or "stochastic terrorism" or "disposable" agents (i.e., individuals who have been recruited or coerced online into carrying out acts of violence or sabotage without an awareness of, or alignment with, the

---

[68] World Economic Forum (2025)
[69] Mosolova (2026)

objectives of the recruiters). These events will require closer integration between policing, intelligence, cybersecurity, and public safety and national security partners. While operational integration will likely benefit from AI-enabled data fusion platforms, it will still depend on human judgement as part of the final packaging and review of intelligence products for consumption by senior decision makers.

- **Insider threats.** Police services must also contend with the risk of insider threats linked to ideological extremism. AI-enabled online environments can facilitate covert radicalization and recruitment of both civilian employees and sworn members, particularly by exploiting grievance narratives found on social media, or by harvesting and exploiting open-source information posted by individuals, in order to distort their thoughts and influence their behaviour. When such dynamics take hold inside an agency, the consequences may extend beyond security breaches to include erosion of professional norms, institutional values, ethical standards, and public trust.

  Mitigating this risk requires sustained leadership attention to workforce integrity, workforce training about online influence and radicalization pathways, early detection (e.g., screening prospective recruits), and organizational cultures (e.g., clear codes of conduct and practice norms, internal reporting mechanisms) that reinforce democratic policing values and that balance freedom of expression, and due process with early intervention.

Scenario | Implications

Scenario: In 2027, a mid-sized Canadian city experiences a cascade of seemingly unrelated events. A business owner is defrauded via a convincing deepfake call impersonating the company's CFO. Seniors report extortion using AI-generated intimate images. A wave of synthetic social media accounts spreads inflammatory allegations of police misconduct, with AI tools amplifying the content to specific neighbourhoods and demographics. Online forums begin calling for "direct action," and a young man - radicalized through encrypted channels by anonymous actors he never met - attempts an act of sabotage at a municipal facility. An internal review simultaneously reveals that a uniformed member had been drawn into extremist online communities through algorithmically curated grievance narratives over several months.

Implications: This is not a set of isolated incidents, it is an ecosystem event. AI has simultaneously accelerated fraud, amplified disinformation, enabled exploitation, and blurred the line between cybercrime and national security.

AI-enabled harms will not arrive neatly categorized. They will converge. Responding effectively will require integrated intelligence, close coordination with cybersecurity and national security partners, and sustained leadership attention to workforce integrity and organizational culture.

Police leaders should assume AI-enabled threats will outpace traditional responses unless adaptive countermeasures are developed and updated at the speed of cyber. That capacity depends as much on organizational culture and leadership as it does on technology. Moreover,

positive community relations and trust will be central to withstanding malicious information operations targeting police-community relations.

Conclusions

> *"Human thinking is messy and full of errors, but it forces us to debate, to doubt, to test ideas against one another – and to recognise the limits of our own understanding. It builds confidence, both individually and collectively.*
>
> *…the exercise of reason was never just about knowledge; it was about enabling people to become agents of their own lives, and resist domination. It was about building a moral community grounded in the shared principle of reason and debate, rather than blind belief.*
>
> *With all the benefits AI brings, the challenge is this: how can we harness its promise … without eroding human reasoning, the cornerstone of the Enlightenment and of liberal democracy itself? That may be one of the defining questions of the 21st century. It is one we would do well not to delegate to the machine."*
>
> - *Joseph de Weck[70], Foreign Policy Research Institute*

Some ongoing AI adoption in policing is unavoidable. Organized crime, fraud networks, and adversarial actors are already leveraging AI-enabled tools. As one interviewee put it, "… police are going to have to engage with AI, even if they only do so defensively, because organized crime is going to use it." Mounting pressures for increased efficiencies driven by labour demands, recruitment and retention challenges, and evolving public expectations will only intensify. Emerging AI-based technologies can enhance organizational and operational capacity and alleviate critical workload pressures.

But AI adoption is not a neutral modernization step. AI systems do not merely improve technical efficiency; they reshape what is visible, prioritized and acted upon. A critical distinction exists between AI as an ***administrative efficiency tool*** vs AI as a ***decision support system*** in operations, intelligence, and risk assessment. These are different risk categories and must not be treated the same. As discussed earlier, a human in the loop does not by itself guarantee meaningful control, particularly for tools such as facial recognition that can anchor a reviewer's judgement under time pressure.

The strategic task for police leaders and boards is to ensure that AI strengthens effectiveness without eroding fundamental rights and institutional legitimacy. This means treating AI as a governed capability, embedded within professional judgement, constrained by law, and accountable to the community. Governance discussions should not just focus on what AI can do, but on what decisions the organization is prepared to allow AI to influence.

Without proactive governance, agencies risk reacting to what communities or employees define as scandals. AI governance is not a communications issue; it is an institutional infrastructure. Successful AI adoption depends as much on people and organizational culture as it does on

[70] De Weck (2025)

technology. Adoption decisions must ensure that individual or cumulative AI systems do not contribute to:

- A gradual erosion of trust both inside and outside the organization;
- A loss of critical thinking, professional discernment and practice skills;
- Disproportionate impact on racialized and other equity-deserving communities;
- Normalized surveillance practices;
- Downstream evidentiary risks in court; and
- Vendor dependence and data sovereignty risks.

These harms may not show up as dramatic single failures. They tend to accumulate, producing system-level damage related to legitimacy, performance and independence. That is difficult to reverse. AI governance must therefore assess not only episodic risk but aggregate impacts on organizational effectiveness, individual dignity and belonging, and institutional trust over time. The decision not to adopt, or to withdraw, a particular AI system should always be on the table, supported by predefined conditions so that the decision does not depend on a crisis to trigger it.

Governance and Oversight Are Essential

The adoption and use of AI systems can compound pre-existing governance gaps, and not every police agency has the capacity required to achieve effective AI governance on its own (see discussion of organizational capture and sector-wide capacity gaps in Section 4 above). This is particularly the case with small and mid-sized agencies, and for services whose technology functions sit within shared municipal environments.

Labour relations are an often-under-examined dimension of this challenge. Discussions about AI in policing tend to give little attention to how these technologies affect employees, including through workplace monitoring, impacts on collective agreements, and the erosion of professional judgement and discretion.

Governance is the most consequential thing police leaders and boards can strengthen right now. In the wake of the Government of Canada's new AI strategy[71], boards and executives will be well-served by building their capacity to assess both the downstream implications of AI adoption and the upstream design decisions and business objectives of AI vendors. This will be especially important as multi-agent systems become more common in business settings, and where the negative impacts of agent misalignment may be significant. That work must also be realistic. A comprehensive framework that stalls under its own weight provides less protection than a smaller set of controls that is actually implemented.

For these reasons, the following recommendations are phased. Five (5) are designated as a Year-1 Core: a governance anchor with a named executive lead and tool inventory; an acceptable-use policy for shadow AI; a tiered risk classification with scope-change triggers; standard contract terms; and AI literacy backed by safeguards against automation bias. These

[71] Innovation, Science and Economic Development Canada (2026)

controls are achievable for services of any size, and they form the foundation on which the remaining recommendations can be built as capacity matures.

In closing, we echo a question posed by Massey College Principal, Dr. James Orbinski, at the opening of a recent college session on: *"Where might AI be taking us, or where might we take AI?".*[72] That question is as pressing for Canadian police agencies as it is for our country.

Recommendations

The following recommendations, grouped by theme, are intended to guide the development of a practical approach to AI governance for Canadian police services. They are phased. The five (5) recommendations marked “Year-1 Core” establish the foundational controls every service should have in place within the first year, regardless of size. The remainder build on that foundation and can be sequenced to each service’s capacity, with priority given to those that apply to medium- and high-risk tools.

Governance and Oversight

Recommendation #1 (Year-1 Core): The chief and board should designate AI adoption and its use as a standing item on the executive governance agenda, with a named executive accountable for AI governance, a maintained inventory of all AI tools in use, and quarterly reporting on deployments, incidents, and review findings, given that faulty or rights-contravening applications carry direct reputational and legitimacy risk to the organization.

Recommendation #2: The board should establish a standing, external review panel - comprising a technical expert, a human rights or privacy specialist, and a community advocate - with the mandate to review all high-risk AI deployments annually and report findings publicly. Except under exceptional circumstances, the panel should have confidential access to method-level information that cannot be disclosed publicly without compromising investigations.

Recommendation #3: Legal counsel (ideally specialized external counsel) should review investigation protocols annually against current case law and emerging standards on AI-related evidence integrity and Crown disclosure obligations.

Recommendation #4: Procurement should include a standard contract clause requiring annual independent performance, bias and social compliance audits for every AI vendor contract, including re-validation following any material update, with audit reports delivered to the board, starting with the next contract renewal cycle.

[72] Orbinski (2026)

Decision Gates

Recommendation #5 (Year-1 Core): The chief, with the police board, should - as a matter of priority - adopt a tiered risk-classification scheme for AI tools (e.g., low/medium/high intrusiveness) that distinguishes administrative uses from operational uses that directly affect the public, and require board approval before any high-risk tool proceeds to procurement. Particularly consequential uses involve applications of AI for decision support. Each approval should define the tool's permitted scope; any material change in use, data sources, or functionality should trigger a new risk assessment and renewed approval.

Recommendation #6: Legal counsel, the privacy officer, and IT security should complete a documented joint sign-off - covering legal authority, privacy impact, cybersecurity risk, and ethical considerations - before any AI tool is deployed operationally, with no exceptions absent board-level emergency authorization. In the private sector, external ethics boards are often used, except in highly sensitive situations. Where IT or procurement functions are delivered through shared municipal services, the sign-off process should extend to those partners through formal agreements. This process may consider findings from reports of the external review panel (Recommendation #2).

Recommendation #7: The chief should direct staff to identify and join (or, absent one, help establish) a provincial or national vetted-tools registry within 12 months, so that future evaluations can draw on other agencies' due diligence rather than starting from zero each time. The registry should be developed in step with the shared evaluation body described in Recommendation #20.

Recommendation #8: The board should define, in advance, the conditions under which an AI tool must be rejected, restricted, or withdrawn (the "zero-option"), including failed audits, unmitigated bias, unauthorized scope expansion, material vendor non-compliance, or a significant loss of public confidence. The board must retain the option to reject any tool at any stage of the process, regardless of prior investment or vendor commitments, and should publish a register of tools in use and tools withdrawn.

Procurement and Risk Management

Recommendation #9: Procurement should require a documented problem statement - including a "safety for whom?" assessment - before any AI acquisition proceeds, and add cybersecurity, inspectability, explainability, data sovereignty, and local control as scored criteria in every AI vendor RFP, weighted explicitly alongside cost and performance, starting with the next procurement cycle.

Recommendation #10 (Year-1 Core): Legal counsel should develop a standard contract template identifying: explicit data-ownership terms; minimum cybersecurity standards; disclosure of system logic and training data provenance sufficient to meet court-ordered and Crown disclosure obligations; vendor provision of expert testimony when AI-derived evidence is challenged; advance notice and re-validation before any update is deployed operationally; protections in the event of a change in vendor ownership; controls on cost escalation; and a

defined exit/data-portability clause. The procurement unit should seek to apply it to every new and renewed AI vendor agreement going forward (recognizing that large foreign vendors may refuse these terms).

Recommendation #11 (Year-1 Core): IT security should update the acceptable-use policy within 90 days to explicitly identify unapproved AI tools (e.g., personal ChatGPT, Claude, or Gemini accounts) as prohibited shadow AI, identify approved and secure alternatives, and incorporate the revised policy into onboarding and annual training.

Recommendation #12: Technology contracts should include provisions for vendors to deliver comprehensive knowledge transfer, including documented procedures, hands-on training, and competency development for agency personnel, to ensure sustainable in-house capability and avoid unnecessary vendor dependency.

Recommendation #13: Where IT, data, or procurement functions are delivered through shared municipal services, the chief and board should establish formal service level agreements and a cross-governance framework with the responsible municipal bodies, ensuring that the police service retains control over Charter, privacy, and evidentiary compliance for any AI tool used in policing.

Workforce and Organizational Preparedness

Recommendation #14 (Year-1 Core): Training and professional development should deliver mandatory AI-literacy training to sworn and civilian employees as a priority within 12 months, and refreshed annually or as tools change, including explicit instruction on automation bias. Training should be paired with structural safeguards: a documented rationale whenever an employee accepts or overrides a consequential AI output, and supervisory audits that confirm primary source material continues to be reviewed. Training should be embedded within systems of supervision, support and performance monitoring that have been updated to ensure effective and appropriate use of AI technologies. In some cases, it may be advantageous for agencies to pool resources to commission or procure joint training packages from a trusted, knowledgeable, supplier.

Recommendation #15: The chief should sequence high-risk AI deployments to internal capacity, deploying no new updates to a high-risk tool until the prior one has completed its first independent review. If several high-risk tools are being simultaneously proposed for adoption, a plan should be in place for all of them to be evaluated at the earliest opportunity. Lower-risk administrative tools may proceed on a faster track, consistent with the risk classification in Recommendation #5. The board retains explicit authority to pause a specific deployment if review findings, public concern, published independent third-party investigations/research or vendor conduct warrant it.

Recommendation #16: Professional standards branches should add online radicalization indicators to existing insider-threat screening and incorporate them into recruit vetting and periodic employee review, in consultation with the association(s)/union(s) on scope and process.

Transparency and Accountability

Recommendation #17: Legal counsel and corporate communications should jointly produce a plain-language explanation template for each deployed AI tool, covering what it does, what data it uses, and how a person can challenge an output, for use with the public and in disclosure to the courts. These explanations should be published through a public register of AI tools in use and tools withdrawn (Recommendation 8), paired with an accessible complaint channel. Any exception for exigent circumstances should be narrowly defined, approved at board level, and be subject to a timeframe for eventual disclosure when the tool is no longer required.

Stakeholder Engagement

Recommendation #18: Tools classified as medium or high risk should undergo a structured community consultation. This would precede board approval of tools identified as being high risk because consultations may result in some medium risk tools being designated as high risk (per Recommendation #5).

> Consultation should be treated as a non-waivable operational requirement and tracked accordingly in the approval record. It should prioritize the communities most likely to be affected, including internal personnel, racialized, marginalized, and Indigenous communities; explain the rationale for the proposed use of the tool; address fear of surveillance as well as demonstrated harms; and carry meaningful decision weight, with the board publishing how community input shaped its decision.
>
> Those participating in the consultation process should be able to engage in substantive dialogue about the proposed tool(s). Community concerns should be able to trigger the external review in Recommendation #2 or the zero-option conditions in Recommendation #8. Any exception for exigent circumstances should be narrowly defined, approved at board level, and followed by consultation after the fact.

Recommendation #19: The board should allocate a specific budget line and executive lead for community engagement within the AI governance framework, and review adequacy of that resourcing annually.

Fragmented Capacity

Recommendation #20: Given the capacity gap documented in Section 4, chiefs should advocate for a shared regional or provincial AI evaluation body at the next relevant chiefs' association meeting, with a goal of establishing a working group within six months and securing support for a feasibility study within 12 months. The body should consolidate and share assessment information on AI systems used in policing, maintain the vetted-tools registry described in Recommendation #7, and give smaller services access to shared technical, legal, privacy, and academic expertise.

Recommendation #21: Consistent with Perrin's (2026) proposal to build on the existing multi-layered oversight ecosystem discussed in Section 4, the board should direct staff to explore and pilot a relevant impact assessment tool on the next medium- or higher-risk AI procurement, and

report results back to the board to inform whether it should become standard practice for all high-risk tools.

> Note: Examples include the Law Commission of Ontario's (2024) *AI Human Rights Impact Assessment Tool* and the Treasury Board of Canada's (n.d.) *Algorithmic Impact Assessment*.

Impact of Recommendations

While technological innovations in AI are accelerating, organizational readiness is generally lagging across most Canadian sectors.[73] Gaps in understanding and trust continue to be important barriers to grappling with the changes that are currently taking place, and for scaling the most promising opportunities.[74] If implemented, these recommendations change how decisions about AI in policing get made. In many services, adoption may be influenced by familiarity with vendors and may also be based on concerns of missing out on new ways to manage pressing operational concerns. It is usually handled one platform or tool at a time, and is typically reviewed only after problems surface.

Under the approach proposed here, every AI tool starts with a clearly defined problem and is matched to its level of risk. It is approved within a defined scope and reviewed again whenever that scope changes. Boards know what platforms and tools are in use, what they are used for, and when to say no. Chiefs have contracts that protect evidence in court, data ownership and the ability to exit. Frontline staff use AI as an aid to their judgement, not a substitute for it, with simple checks that keep the original evidence in view.

The effects reach beyond the organization:

- Communities most affected by AI are involved and have influence over how it is used, and a clear way to raise concerns.
- Courts and oversight bodies get the disclosure they need to test AI-derived evidence.
- Smaller agencies no longer have to assess complex technology alone. They can draw on shared evaluation, vetted tools and pooled expertise, which reduces duplicated effort and repeated mistakes across the sector.
- Relationships of trust between police and community – contributors to democracy and sovereignty – are strengthened.

Most importantly, the proposed approach protects public trust while there is still time to protect it. The harms this report identifies tend to build up quietly and are hard to reverse.[75] The potential benefits of AI systems will be better realized and sustained through carefully planned adoption strategies. Acting now, starting with the five (5) Year-1 core measures, lets police services gain the operational benefits of AI without trading away legitimacy, independence or

---

[73] Georgian (2026)
[74] Gillespie, et al. (2025)
[75] cf. Kasirzadeh (2025) on "accumulative" AI risk

the professional judgement policing depends on. Leadership on this must come from police services, not from vendors or as a result of reputational damage after a crisis has occurred. A provisional set of decision guideposts for AI adoption is offered in Appendix A.

Areas Warranting Further Study

This report provides foundational guidance on data governance. However the intersection of AI and Indigenous data sovereignty presents distinct and complex challenges that require further investigation. In addition, traditional responsibility centres related to governing the adoption and use of IT systems may no longer be sufficient with the expansion of AI.

Indigenous Data Sovereignty in AI Frameworks

The rapid development and integration of AI tools in Canadian policing introduce substantial complexities regarding data management, particularly concerning data collected from, or generated about, Indigenous people and communities. While frameworks like the **OCAP® principles** (Ownership, Control, Access, and Possession) for First Nations[76], or those touching on the concerns and aspirations of Métis and Inuit peoples[77] offer a robust starting point for data governance, standardizing this remains an unaddressed challenge.

To properly address these issues, further research is required in the following areas:

- Investigate and understand how data sovereignty principles uniquely apply to, and differ between, First Nation, Métis, and Inuit communities, avoiding a singular approach to Indigenous data.
- Consider mechanisms for ethical data governance and partnership models tailored specifically for the significant populations of Indigenous Canadians living in urban areas, who may fall outside traditional community-based governance structures.
- Research practical technical pathways to embed data sovereignty rules directly into AI data architecture, focusing on such aspects as cloud storage decisions, machine learning model training datasets, and access control protocols.
- Identify appropriate approaches for establishing meaningful co-design partnerships with relevant Indigenous organizations and rights-holders to ensure compliance with emerging Indigenous data rights.

Anchoring AI Accountability at the Executive Level

Recognizing the growing demands that AI is placing on governance, some management scholars have suggested that several key functions could be anchored to the role of chief compliance officer (CCO).[78] The role would be empowered with the authority to influence decisions, independent of the units that use specific AI tools. It could also consolidate and

[76] First Nations Information Governance Centre (2019; 2021)
[77] Inuit Tapiriit Kanatami (2018); Métis National Council (2025)
[78] Oxéus, et al. (2026)

coordinate decisions that currently tend to be spread across legal, privacy, IT security, procurement and professional standards.

Oxéus, et al. (2026) suggest that this position should sit at the executive table and report directly to the board, as chief information security officers (CISOs) in the private sector increasingly do. For smaller services, this capacity may be best provided through the shared regional evaluation body proposed in Recommendation #20.

Although such an approach suggests a promising avenue for organizational capacity building, it is largely conceptual and so remains to be assessed as a practical way of dealing with the aforementioned issues within police agencies.

References

Adams, I.T. (2025). Automation and artificial intelligence in police body-worn cameras: Experimental evidence of impact on perceptions of fairness among officers. *Journal of Criminal Justice*, *97*, Article 102330. https://www.sciencedirect.com/science/article/pii/S0047235225000224

Angus Reid Institute. (2025, November 20). AI use is widespread, but concern is rising among Canadians. https://angusreid.org/artificial-intelligence-ai-use-concerns/

Angus Reid Institute. (2026, June 1). Canadians call for heavy AI regulation, but three-quarters doubt any government can keep up with the technology. https://angusreid.org/ai-regulation-canada/

Angus Reid Institute. (2026, June 8). AI at work: Few Canadians see impact as significantly positive as many worry over downsides of increased use. https://angusreid.org/ai-at-work/

Bachar, C. (2020, September 23). Report on body-worn cameras complete. The Oshawa Express. https://oshawaexpress.ca/report-on-body-worn-cameras-complete/

Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation*. Thousand Oaks, CA, USA: Sage.

Battagello, D. (2020, March 4). Clearview AI facial recognition worrisome for Black, Indigenous, marginalized communities: Experts. CBC News. https://www.cbc.ca/news/canada/windsor/clearview-ai-worrisome-for-black-indigenous-marginalized-communities-1.5486325

Bender, E. & Hanna, A. (2025). *The AI con*. New York, NY: Harper-Collins.

Booth, R. (2026, February 22). Met police using AI tools supplied by Palantir to flag officer misconduct. *The Guardian*. https://www.theguardian.com/uk-news/2026/feb/22/met-police-ai-tools-officer-misconduct-palantir

Bhuiyan, J. (2026, August 18). OpenAI announces slowing pace of development after hack by rogue agent. *The Guardian*. https://www.theguardian.com/technology/2026/aug/18/open-ai-pause-hack

Botschner, J., Cioffi, G., McMahon, D., Ollinger, J., Sylvestre, B., Kotak, R. & Corley, C. (2025, July 30). Understanding and countering gendered disinformation: A framework for resilience and action. Ottawa, ON: Community Safety Knowledge Alliance. https://cskacanada.ca/wp-content/uploads/2025/07/R_Understanding-and-Countering-Gendered-Disinformation-MAIN-REPORT-Annexes-July-30-2025.pdf

Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review*, 82(5), 997-1008. https://doi.org/10.1177/0003122417725865

Canada. (1982). Canadian Charter of Rights and Freedoms (Part I of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (UK), 1982, c. 11). Justice Laws Website. https://laws-lois.justice.gc.ca/eng/const/page-12.html

Canada. (1985). *Privacy Act*, R.S.C. 1985, c. P-21. Justice Laws Website. https://laws-lois.justice.gc.ca/eng/acts/P-21/

Canada. (2000). *Personal Information Protection and Electronic Documents Act*, S.C. 2000, c. 5. Justice Laws Website. https://laws-lois.justice.gc.ca/eng/acts/P-8.6/

Canadian Centre for Cyber Security. (2024, October 30). National cyber threat assessment 2025–2026. Communications Security Establishment. https://www.cyber.gc.ca/en/guidance/national-cyber-threat-assessment-2025-2026

Canadian Civil Liberties Association. (2022). Facial recognition technology and policing in Canada. https://ccla.org

Carleton, C. & Guirado, J. (2026). What's at stake: Preserving what makes us deeply human in the age of AI. Center for Humane Technology, Substack, February 1, 2026. Retrieved February 3, 2026. https://centerforhumanetechnology.substack.com/p/whats-at-stake-preserving-what-makes

Castaldo, J. & Posadzki, A. (2026, October 13). Scientists push for creation of international body to regulate artificial intelligence. *The Globe and Mail*. https://www.theglobeandmail.com/business/article-scientists-international-body-regulate-artificial-intelligence/

CBC News. (2025, June 11). What you need to know about the proposed AI data centre in Hamilton. CBC News. https://www.cbc.ca/news/canada/hamilton-data-centre-slate-9.7223450

Cecco, L. (2026, October 13). AI godfather Geoffrey Hinton: 'We're creating alien beings. I'm willing to be selfish and choose humans'. *The Guardian*. https://www.theguardian.com/us-news/2026/oct/13/geoffrey-hinton-trump-ai-tech

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In (N. K. Denzin and Y. S. Lincoln, Eds) *Handbook of Qualitative Research* (2nd Ed.). Thousand Oaks, CA, USA: Sage, pp. 509–535.

Conger, K., Isaac, M., Barnes, J. E., & Sanger, D. E. (2026, October 9). Anthropic agents tried to fill out visa forms on State Dept. website. *The New York Times*. https://www.nytimes.com/2026/10/09/technology/anthropic-rogue-ai-agents.html

Crawford, K. (2021). *Atlas of AI*. New Haven, CT: Yale University Press.

Crootof, R., Kaminski, M. E., & Price, W. N., II. (2023). Humans in the loop. *Vanderbilt Law Review*, 76(2), 429–519. https://scholarship.law.vanderbilt.edu/vlr/vol76/iss2/2

Deloitte Services LP. (2025, April 30). Generative AI risks and how to manage them. *The Wall Street Journal* (Deloitte CIO Journal). https://deloitte.wsj.com/cio/generative-ai-risks-and-how-to-manage-them-75580c9a

de Weck, J. (2025, December 26). Our king, our priest, our feudal lord — how AI is taking us back to the dark ages. *The Guardian*. https://www.theguardian.com/commentisfree/2025/dec/26/ai-dark-ages-enlightenment?CMP=Share_iOSApp_Other

Doctorow, C. (2026). *The reverse centaur's guide to life after AI: How to think about artificial intelligence – before it's too late*. New York, NY: Farrar, Straus and Giroux.

Du Sault, L. (2023, July 9). Under union pressure, Vallejo police chief ends body camera analysis. *Open Vallejo*. https://openvallejo.org/2023/07/09/under-union-pressure-vallejo-police-chief-ends-body-camera-analysis/

Fanfair, R. (2026, April 23). Unprecedented SMS blaster arrests. Toronto Police Service. https://www.tps.ca/media-centre/stories/unprecedented-sms-blaster-arrests/

First Nations Information Governance Centre. (2019). First Nations data sovereignty in Canada. *Statistical Journal of the IAOS*, 35(1), 47–69. DOI 10.3233/SJI-180478

First Nations Information Governance Centre. (2021). The First Nations principles of OCAP®. First Nations Information Governance Centre. https://fnigc.ca/ocap-training/

Georgian. (2026). *Canada doubled its AI runners in a year but fell behind on agentic AI* [Benchmark report]. https://georgian.io/benchmark-reports/canadian-ai-runners

Gillespie, N., Lockey, S., Ward, T., Macdade, A., & Hassed, G. (2025). Trust, attitudes and use of artificial intelligence: A global study 2025. The University of Melbourne and KPMG. DOI 10.26188/28822919.

Globe and Mail Editorial Board. (2026, February 27). Don't leave AI safety to the tech bros. *The Globe and Mail*. https://www.theglobeandmail.com/opinion/editorials/article-ai-regulation-tumbler-ridge/

Government of Canada, Ministry of Innovation, Science and Industry, & Ministry of Public Safety. (2022, May 19). Government of Canada announces intention to ban Huawei and ZTE from Canada's 5G networks [Policy statement]. https://www.canada.ca/en/innovation-science-economic-development/news/2022/05/statement-from-minister-champagne-on-telecommunications-security.html

Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *BMJ*, 331(7524), 1064–1065. https://doi.org/10.1136/bmj.331.7524.1064

Guan, J., Blanchard, T., Foerster, H., Jia, H., Huang, G., & Papernot, N. (2026). *AI agents enable adaptive computer worms* [Preprint]. arXiv. https://arxiv.org/abs/2606.03811

Hahmann, T., & Lovei, M. (2026). *Workplace artificial intelligence use: A profile of sociodemographic and job characteristics* (Insights on Canadian Society, Catalogue No. 75-006-X). Statistics Canada. https://www150.statcan.gc.ca/n1/pub/75-006-x/2026001/article/00007-eng.htm

Huey, L., Ferguson, L., & Towns, Z. (2025). “Cops need doxxed”: Releasing personal information of police officers as a tool of political harassment. *Crime & Delinquency*, *71*(3), 714–739. https://doi.org/10.1177/00111287241271069

Inuit Tapiriit Kanatami. (2018). *National Inuit strategy on research*. https://www.itk.ca/wp-content/uploads/2020/10/ITK-National-Inuit-Strategy-on-Research.pdf

Innovation, Science and Economic Development Canada. (2026, June 4). Canada's national artificial intelligence strategy: AI for all. Government of Canada. https://ised-isde.canada.ca/site/ised/en/canadas-national-artificial-intelligence-strategy-ai-all

International Association of Chiefs of Police (2026). IACP Connector. Washington, DC. https://www.theiacp.org/resources/iacp-connector

Internet Watch Foundation. (2026). Harm without limits: AI child sexual abuse material through the eyes of our analysts. Internet Watch Foundation. https://admin.iwf.org.uk/media/hl1nvdti/iwf_ai-CSAM-report-2026.pdf

Jacobs, G., van Houdt, F., & coons, g. (2024). Studying surveillance AI-cologies in public safety: How AI is in the world and the world in AI. *Surveillance & Society*, 22(2), 160–178. https://doi.org/10.24908/ss.v22i2.16104

Kasirzadeh, A. (2025). Two types of AI existential risk: Decisive and accumulative. *Philosophical Studies*, *182*(7), 1975–2003. https://doi.org/10.1007/s11098-025-02301-3

Kaye, D. (2026, February 27). Burger King AI bot will check up on staffs' please and thank yous. BBC News. https://www.bbc.com/news/articles/cgk2zygg0k3o

Ko v. Li, 2025 ONSC 2965 (CanLII). https://canlii.ca/t/kc6xx

Law Commission of Ontario, & Ontario Human Rights Commission. (2024). Human rights AI impact assessment (HRIA): Guide. https://lco-cdo.org/en/publications/human-rights-ai-impact-assessment-hria/

Lawrence, M., Homer-Dixon, T., Janzwood, S., Rockström, J., Renn, O., & Donges, J. F. (2024). Global polycrisis: The causal mechanisms of crisis entanglement. *Global Sustainability*, 7, Article e6, 1–16. https://doi.org/10.1017/sus.2024.1

Lockwood, C., Dobson, C., Phillips, C., & Pedersen, R. (2025, July 14). Artificial advocacy: How Canadian courts and legislators are responding to generative AI. Osler, Hoskin & Harcourt LLP. https://www.osler.com/en/insights/updates/artificial-advocacy-how-canadian-courts-and-legislators-are-esponding-to-generative-ai/

Mehdi, T., & Morissette, R. (2024). *Exposure to artificial intelligence in Canadian jobs: Experimental estimates* (Analytical Studies Branch Research Paper Series, Catalogue No. 11F0019M). Statistics Canada. https://.www150.statcan.gc.ca/n1/pub/11f0019m2024005-eng.htm

Métis National Council. (2025). *Métis Nation data strategy*. https://www.metisnation.ca/wp-content/uploads/2025/12/Consolidated_Me_tis_Nation_Data_Strategy_Final_15_10_2025_3_56.pdf

Milmo, D., & Down, A. (2026, June 3). Future AI weapons such as drones should have moral code, says former UK spy chief. *The Guardian*. https://www.theguardian.com/science/2026/jun/03/ai-weapons-drones-moral-code-former-uk-gchq-chief-david-omand

Moser, A., & Korstjens, I. (2017). Series: Practical guidance to qualitative research. Part 1: Introduction. *European Journal of General Practice*, *23*(1), 271–273. https://doi.org/10.1080/13814788.2017.1375093

Mosolova, D. (2026, February 27). How AI is supercharging Russia's online disinformation campaigns. BBC News. https://www.bbc.com/news/articles/cx2r7grrdwzo

MPS confirms the use of AI to spot potential misconduct. (2026, February 23). *Police Professional*. https://policeprofessional.com/?p=143282

Noguchi, Y. (2023, June 8). An eating disorders chatbot offered dieting advice, raising fears about AI in health. NPR. https://www.npr.org/sections/health-shots/2023/06/08/1180838096/an-eating-disorders-chatbot-offered-dieting-advice-raising-fears-about-ai-in-hea

Office of the Privacy Commissioner of Canada. (2021). Guidance on the use of artificial intelligence by federal institutions. https://www.priv.gc.ca

Office of the Privacy Commissioner of Canada. (2021). PIPEDA Findings #2021-001: Joint investigation of Clearview AI, Inc. by the Office of the Privacy Commissioner of Canada, the Commission d'accès à l'information du Québec, the Information and Privacy Commissioner for British Columbia, and the Information Privacy Commissioner of Alberta. https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2021/pipeda-2021-001/

Office of the United Nations High Commissioner for Human Rights (OHCHR), & Asia Pacific Forum of National Human Rights Institutions (APF). (2013). The United Nations Declaration on the Rights of Indigenous Peoples: A manual for national human rights institutions. United Nations. https://www.ohchr.org/sites/default/files/Documents/Issues/IPeoples/UNDRIPManualForNHRIs.pdf

Ontario Human Rights Commission, & Law Commission of Ontario. (2024). Human rights AI impact assessment. Ontario Human Rights Commission.

OpenAI. (2026). *OpenAI–Hugging Face incident technical report* [Technical report]. https://cdn.openai.com/pdf/67869394-cb91-4c12-888c-5cbd85c7814c/OpenAI-Hugging-Face%20Incident-Technical-Report.pdf

Orbinski, J. (2026, February 26). Opening remarks. (Eh): Artificial intelligence in the Canadian context. Massey College AI and Human Societies panel discussion. Toronto, ON: Massey College.

Osler, Hoskin & Harcourt LLP (2026). *The Protecting Privacy and Consumer Data Act (Bill C-36): Key obligations and enforcement review.* June 2026. https://www.osler.com/wp-content/uploads/2026/07/Guide-to-the-PPCDA-Bill-C-36.pdf

Oxéus, G., Koplyay, T., & Gagnon, S. (2026). Chief Compliance Officers Navigating Ethical, Integrity, and Responsibility Challenges in Artificial Intelligence (AI) Projects and against Digital Disinformation. In *The Changing Role of the Chief Compliance Officer* (pp. 106-140). Routledge.

Parliament of Canada (2026). *Bill C-36: An Act to enact the Protecting Privacy and Consumer Data Act, to amend the Personal Information Protection and Electronic Documents Act and to make amendments to other Acts*. https://www.parl.ca/legisinfo/en/bill/45-1/c-36

Pawson, R., Boaz, A., Grayson, L., Long, A., & Barnes, C. (2003). *Types and quality of knowledge in social care* (Knowledge Review No. 3). Social Care Institute for Excellence. https://lx.iriss.org.uk/sites/default/files/resources/kr03.pdf

Perrin, B. (2026). AI & policing: Roundtable report on the governance & use of artificial intelligence by police in Canada. UBC AI & Criminal Justice Initiative. https://canlii.ca/t/7nr4b

Perrin, B., Liew, G., & Sweeney, I. (2026). AI & policing: Research report on the governance & use of artificial intelligence by police in Canada (2026 CanLIIDocs 293). UBC AI & Criminal Justice Initiative. https://canlii.ca/t/7nrqc

R. v. Mohan, [1994] 2 S.C.R. 9. CanLII. https://www.canlii.org/en/ca/scc/doc/1994/1994canlii80/1994canlii80.html

R. v. Stinchcombe, [1991] 3 S.C.R. 326. CanLII. https://www.canlii.org/en/ca/scc/doc/1991/1991canlii45/1991canlii45.html

Rose, S. (2026, February 25). Tech legend Stewart Brand on Musk, Bezos and his extraordinary life: 'We don't need to passively accept our fate'. *The Guardian*. https://www.theGuardian.com/technology/2026/Feb/25/tech-legend-stewart-brand-on-musk-bezos-and-his-extraordinary-life-we-dont-need-to-passively-accept-our-fate?cmp=share_iosapp_other

Royal Canadian Mounted Police. (2020). Artificial intelligence and policing: Considerations for law enforcement. https://www.rcmp-grc.gc.ca

Sims, D. (2026, February 7). *Waymo admits that its autopilot is often just guys from the Philippines*. TechSpot. https://www.techspot.com/news/111233-waymo-admits-autopilot-often-guys-philippines.html

Standing Committee of the National People's Congress of the People's Republic of China. (2017, June 27). National Intelligence Law of the People's Republic of China (DiCarlo, Trans.). China Law Translate. https://www.chinalawtranslate.com/en/national-intelligence-law-of-the-peoples-republic-of-china-draft/

Statistics Canada. (2023). Police resources in Canada. https://www150.statcan.gc.ca

TechCabal. (2026, March 4). Kenyan workers say Meta Ray-Ban AI glasses expose intimate moments. https://techcabal.com/2026/03/04/kenya-meta-ray-ban-ai-intimate-moments/

The Associated Press. (2025, November 6). Denmark eyes new law to protect citizens from AI deepfakes. https://www.ap.org/news-highlights/spotlights/2025/denmark-eyes-new-law-to-protect-citizens-from-ai-deepfakes/

Treasury Board of Canada Secretariat. (2023). Directive on automated decision-making. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592

Treasury Board of Canada Secretariat. (n.d.). Algorithmic impact assessment. Government of Canada. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/automated-decision-making/algorithmic-impact-assessment.html

Truth and Reconciliation Commission of Canada. (2015). Truth and Reconciliation Commission of Canada: Calls to action. http://nctr.ca/assets/reports/Calls_to_Action_English2.pdf

Tuquero, L. (2026, September 2). AI agents are hacking systems without any input from humans. How did we get here? *PBS NewsHour*. https://www.pbs.org/newshour/science/ai-agents-are-hacking-systems-without-any-input-from-humans-how-did-we-get-here

United Nations Children's Fund. (2025). Artificial intelligence and child sexual abuse and exploitation: Emerging risks and policy considerations. UNICEF. https://www.unicef.org/media/178571/file/UNICEF%20AI%20CSEA%20Brief_2.pdf

United Nations Office on Drugs and Crime. (2025). Emerging threats: The intersection of criminal and technological innovation in the use of automation and AI. United Nations Office on Drugs and Crime. https://www.unodc.org/roseap/uploads/documents/Publications/2025/UNODC_Report_Emerging_threats_-_The_intersection_of_criminal_and_technological_innovation_in_the_use_of_automation_and_AI.pdf

Watson, S. (2023). Investigating the role of social media abuse in gender-based violence: The experiences of women police officers. *Criminology & Criminal Justice*, *23*(5), 829–844. https://doi.org/10.1177/17488958221087487

Watts, F., del Pozo, B., White, M.D. & Malm, A. (2025). Does automated feedback impact the acceptability of AI-generated police body worn camera review? An implementation science natural experiment. *Journal of Experimental Criminology*. https://doi.org/10.1007/s11292-025-09711-7

Wischmeyer, T. (2019). Artificial intelligence and transparency: Opening the black box. In T. Wischmeyer & T. Rademacher (Eds.), *Regulating artificial intelligence*, Springer Nature (pp. 75-101). https://doi.org/10.1007/978-3-030-32361-5_4

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on Evaluation and*

Assessment in Software Engineering (EASE '14). ACM. https://doi.org/10.1145/2601248.2601268

Wong, W.H. (2023). *We, the data: Human rights in the digital age*. Cambridge, MA: The MIT Press.

World Economic Forum. (2025). The future of jobs report 2025. World Economic Forum. https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf

Acknowledgements

CSKA and the authors sincerely thank the distinguished panel of subject matter experts who generously reviewed and commented on an earlier version of this paper. Their candid insights and practical advice greatly strengthened the final report. Panel members took part in their personal capacities. Responsibility for the content rests with the authors.

We extend particular thanks to:

Alfonso Adamo	York Regional Police Service
Ron Anderson	Government of Alberta
Mara Cairo, P.Eng., PMP	Edmonton Police Service
Gideon Christian, PhD	University of Calgary
Christa Dicks, MBA, MEd	Cape Breton Regional Municipality
Nathan Dokis	Greater Sudbury Police Service
Gabrielle Jacobs, PhD	Erasmus University, School of Social and Behavioural Sciences, The Netherlands
Dale Johnstone	Saskatoon Police Service
Guy Leeson	Delta Police Department
Jamie Reynolds	Cape Breton Regional Police Service
Christiane Sadeler, MA	More Better Solutions
Howard Sapers	Canadian Civil Liberties Association
Trent Stevely	Regina Police Service
Tony Ventura, MA	Peel Regional Police Service

Report Authors

Janos Botschner, PhD

Janos is a social scientist who works at the intersection of issues related to community safety and well-being. He previously held senior positions in the broader public sector and academia, including as founding director of a community-based research institute, and chief researcher at a large branch of the Canadian Mental Health Association, where he was involved in the design and evaluation of policies and programs in the context of mental health transformation in Ontario. He has served on the editorial board of the Canadian Journal of Community Mental Health, and as a section chair of the Canadian Psychological Association.

He maintains an active practice related to the intersection of emerging technologies and community safety and well-being, with particular attention to cybersecurity and to the information environment. Over the past three decades, his mission has been to help build the capacity of communities and society to be safe, healthy, inclusive, supportive, and innovative.

Ritesh Kotak, JD, MBA

Ritesh is a technologist and legal advisor whose career uniquely bridges public safety, emerging technology, and innovation. He spent eight years in Canadian policing, where he co-developed a cybercrime unit and led public safety innovations, before transitioning to Big Tech to drive cybersecurity, IoT, and intelligence projects globally.

Currently practicing law, Ritesh advises clients on complex matters ranging from cyber breaches to corporate litigation. An accomplished subject matter expert, he has twice testified before Canadian Parliamentary Committees on digital competitiveness. His insights are frequently featured in mainstream global media, including *The Economist* and the *Financial Times*, and he has been recognized by IT World Canada as a leading cybersecurity specialist.

Ritesh holds a Juris Doctor from the University of Ottawa, an MBA with distinction from the University of Edinburgh, a Bachelor's of Business Administration from the University of Toronto, and completed the Emerging Leaders Executive Program at Harvard's Kennedy School.

Neil Dubord, PhD

Neil is a distinguished police and business leader with over 30 years of experience. He has served as Chief of the Delta Police Department, Chief of Metro Vancouver Transit Police, and Deputy Chief of the Edmonton Police Service. His exemplary service has earned him numerous honours, including the Queen's Platinum Jubilee Pin, multiple Police Exemplary Service Medals, and Officer of the Police Order of Merit.

A recognized thought leader, Neil has published widely on policing, leadership, and public trust. He holds a Ph.D. in Business Management, a Master's in Leadership and Training, and advanced certifications from the FBI National Academy and University of Virginia. Neil is an active member of the IACP, FBI National Academy, FBI-LEEDA, and the Canadian Association of Chiefs of Police.

Cal Corley, MBA

Cal is the CEO of the Community Safety Knowledge Alliance, a role he has held for the past 11 years. He is a former Assistant Commissioner of the Royal Canadian Mounted Police. From 2008–2014, he was head of the Canadian Police College and during that time served as the RCMP's Senior Envoy to Mexico and the Americas.

Over the course of his career, Cal gained extensive domestic and international experience in both operations and executive management, serving in such areas as community policing, major crime, drug enforcement, national security, criminal intelligence, strategic planning, human resources, and leading major reform initiatives. He also served on secondments at the Privy Council Office's Security and Intelligence Secretariat, and at Public Safety Canada.

About CSKA

The Community Safety Knowledge Alliance (CSKA) is a Canadian non-profit research and education organization dedicated to advancing community safety and well-being across Canada.

Our mission is to conduct independent, evidence-based research on multi-sectoral approaches to improving community safety, and to mobilize that knowledge so it strengthens the delivery and outcomes of community safety and human services for Canadians.

Independence and objectivity are central to CSKA's identity. Our work is built on upholding these values in every relationship with the governments, organizations, and communities we serve, allowing us to sustain a public-interest, evidence-based perspective alongside effective collaboration across sectors.

For more information on CSKA, please visit www.cskacanada.ca or contact CEO Cal Corley at (613) 297-6728 or ccorley@cskacanada.ca.

AI Disclosure

In preparing this report, the authors used Perplexity for discovery, under human direction, to identify relevant literature. Claude was used during copy editing to identify typographical errors and inconsistencies and to flag opportunities to improve clarity.

The analysis, findings, and recommendations are the authors' own and reflect the review by CSKA's multidisciplinary expert panel. AI was not used to generate recommendations. The authors take full responsibility for the accuracy and integrity of this report.

Appendix A: Sample Decision Guideposts for AI Adoption

Governance frameworks often emerge in response to failures, controversies, or unintended consequences, after trust has already been damaged thereby narrowing the options available to decision makers. AI adoption and use of AI technologies in Canadian policing requires a different approach.

The following questions are intended to support anticipatory governance by helping police leaders, boards, and policymakers…

- identify risks and opportunities early
- assess the longer-term implications of AI-enabled technologies
- develop appropriate pathways to adoption, oversight and accountability structures, and

…before they are deployed at scale.

A. Strategic Justification & Necessity

1. Problem Definition

- What operational or community safety problem is this tool intended to solve?
- Is this a capacity problem, a speed problem, an accuracy problem or a judgement problem?
- Could non-AI alternatives achieve similar results?

2. Necessity & Proportionality

- Is AI necessary, or merely convenient?
- Is the scope of deployment proportionate to the stated objective?
- Does the tool expand surveillance beyond the original purpose?

3. CSWB Alignment

- Does this tool strengthen conditions of safety, dignity, and belonging?
- Could it improve efficiency while degrading trust?
- How does this affect upstream prevention vs downstream enforcement?

B. Evidence & Validation

4. Validity

- Has the tool been independently evaluated? If yes, how?
- Is there evidence of accuracy across diverse Canadian populations?
- What are known error rates?

5. Bias & Equity Testing

- Has bias testing been conducted?
- Across which protected groups?
- Who conducted it, and can results be verified?

6. Context Sensitivity

- Is validation based on Canadian legal and social conditions?
- Are foreign benchmarks being inappropriately relied upon?

C. Human Judgement, Professional Discernment & Discretion

7. Human Involvement

- Where precisely does human oversight occur?
- Is the human empowered to override the output?
- Are override decisions documented?

8. Cognitive Bias Mitigation and Practice Discernment

- Have personnel been trained on automation bias and anchoring effects?
- How may the tool(s) be used to enhance employee performance without eroding the capacity for professional discernment and discretion?
- Are safeguards in place to prevent "rubber stamping"?

9. Role Clarity

- Who is accountable if the AI output is wrong?
- Does responsibility shift upward, downward, or diffuse?

D. Legal, Human Rights & Disclosure Implications

10. Charter & Human Rights Review

- Has a formal rights impact assessment been conducted?
- Are privacy, equality, expression, and due process engaged?

11. Mosaic & Cumulative Effects

- Does this tool aggregate or link datasets?
- Could it produce harms only visible at scale?

12. Disclosure & Discoverability

- If used in investigations, is the algorithm's role discoverable?
- Are proprietary restrictions compatible with disclosure obligations?
- Will the vendor disclose system logic and training data, and provide expert testimony, if AI-derived evidence is challenged in court?

E. Governance & Transparency

13. Approval Pathway

- Is this an administrative use, or an operational use that directly affects the public?
- Which uses require board approval?
- Should testing and approval be required upon every update?
- Which require independent review?
- Is there a risk categorization framework in place?
- Are there limitations as to which data sets can and cannot be shared?
- How will the community be meaningfully involved in approval and ongoing governance?

14. Scope Control

- What prevents "function creep"?
- Would expanding the tool's use require renewed approval?
- Is the tool's approved scope documented, and who is responsible for monitoring it?

15. Public Transparency

- What will be publicly disclosed?
- How will communities know when AI is being used?
- Will the tool appear on a public register of AI tools in use and tools withdrawn?

F. Procurement & Vendor Risk

16. Business Requirements

- Have we described our business requirements clearly, and based on input from relevant stakeholders?
- What problem are we solving, and for whom - who benefits, and who bears the risk?
- Does the requirement start from an operational or community need, rather than from a vendor's proposal?

17. Vendor Dependence

- Are we becoming/might we become dependent on a single platform?
- Are exit clauses and data portability protections in place?
- What mitigations can we put in place to lessen dependence and manage risks associated with exit?
- What training and knowledge transfer processes would be available, and at what costs, to ensure that in-house capability can be developed?

18. Platform & Data Sovereignty, Security & Vendor Risk

- Where is the platform/service headquartered?
- Who owns the platform/service?

- Where is the data stored?
- Where is the data processed?
- Who owns the data?
- Can the vendor use it for training?
- Where does data-in-transit travel?
- What vendor cybersecurity measures are in place, including standards and auditing?
- Has the vendor ever been subject to sanctions or legal proceedings?
- Is the vendor associated with any reputational concerns that could impact the public's trust of our agency?
- How will long-term licensing, integration, and support costs be controlled?
- What happens to our data and contract if the vendor is acquired or changes ownership?
- Is the data transferable?
- If First Nations data are involved, how are OCAP® principles applied?

19. Auditability
    - Can the system be independently audited? If yes, how often? Can random audits be conducted?
    - Is the model explainable?
    - Will the vendor notify us, and allow re-validation, before updates are deployed?

G. Monitoring, Evaluation & Exit

20. Performance Monitoring
    - What indicators will track effectiveness and other business requirements?
    - What indicators will track trust and legitimacy?

21. Sunset & Pause Mechanisms
    - What predefined conditions would require deployment to be paused, restricted, or withdrawn?
    - Is there a mandatory review period?
    - What are the contract parameters and terms of exit?
    - Who owns historical data?

22. Data Portability
    - How, and in what form(s), would historical data owned by our agency be transferred to us during exit?

23. Community Feedback
    - Is there a mechanism for affected communities to raise concerns?

- Do affected communities have meaningful influence over the decision, not only an opportunity to comment?
- How will legitimacy damage be detected early?
- Is there a communication plan to proactively disclose certain information such as the usage of the tool to the public?