SAI CLOUD SECURE

WHITEPAPER

AI Security & ISO 42001

Governing Artificial Intelligence in the Public Sector

March 2026 | saicloudsecure.com

About This Whitepaper

This whitepaper examines the growing imperative for AI governance in Canadian federal and provincial government environments. It provides a practical overview of ISO/IEC 42001:2023, the international standard for AI management systems, and offers a roadmap for public sector organizations seeking to adopt AI responsibly, securely, and in alignment with Treasury Board Secretariat expectations.

Intended Audience

CIOs, CISOs, IT Security Architects, and Procurement Officers in Federal and Provincial Government | Enterprise Technology Leaders adopting AI workloads

Contact: kp@saicloudsecure.com | www.saicloudsecure.com | Toronto, Ontario, Canada

Executive Summary

Artificial intelligence is no longer a future-state consideration for the public sector — it is already being deployed across federal departments, Crown corporations, and provincial agencies to automate decisions, process sensitive data, and augment human judgement. Yet the governance frameworks required to manage AI safely, transparently, and accountably have lagged behind the pace of adoption.

ISO/IEC 42001:2023 — the world's first international standard for AI Management Systems (AIMS) — provides a structured, risk-based approach to AI governance that is increasingly relevant to Canadian government procurement, operations, and compliance obligations. For organizations operating under the Treasury Board Secretariat's Directive on Automated Decision-Making, managing Protected B workloads, or handling sensitive citizen data, ISO 42001 offers a credible framework to demonstrate responsible AI stewardship.

This whitepaper explores:

- The unique AI security risks facing public sector organizations
- What ISO/IEC 42001 requires and how it maps to existing government frameworks
- Key implementation steps for federal and provincial departments
- How SAI Cloud Secure helps government clients achieve and sustain AI governance maturity

1. The AI Imperative in Government — and the Governance Gap

1.1 Rapid AI Adoption Across the Public Sector

Canadian government departments are deploying AI and machine learning in areas including fraud detection, case management, natural language processing for citizen services, predictive analytics for resource allocation, and automated document classification. The Treasury Board Secretariat's Directive on Automated Decision-Making (DADM), in effect since 2019 and updated in 2022, mandates impact assessments and transparency requirements for automated decisions affecting citizens — yet enforcement and organizational readiness remain inconsistent.

Key Statistic

A 2025 survey by the Office of the Privacy Commissioner found that 68% of federal institutions reported deploying at least one AI system for internal or citizen-facing decisions, yet fewer than 30% had a formal AI governance policy in place.

1.2 Unique Risks in the Public Sector

AI governance in government is distinctly complex for several reasons:

- Sensitive citizen data at scale — health records, tax information, immigration files, and national security data are routinely processed, amplifying the impact of AI errors or breaches.
- Low tolerance for algorithmic bias — automated decisions affecting benefits, services, or enforcement carry legal, ethical, and reputational consequences.
- Supply chain exposure — government AI systems increasingly rely on third-party models, cloud APIs, and foundation model providers who may themselves have limited transparency into training data provenance.
- Adversarial threats — AI models can be targeted by prompt injection, data poisoning, model inversion, and membership inference attacks — threat vectors not covered by traditional IT security frameworks like ITSG-33.
- Regulatory expectations — beyond DADM, departments face obligations under the Privacy Act, PIPEDA successor legislation (Bill C-27), and pending AI-specific regulation under the proposed Artificial Intelligence and Data Act (AIDA).

1.3 The Governance Gap

Most government IT security programs are anchored in frameworks such as ITSG-33 (IT Security Risk Management), PBMM (Protected B, Medium Integrity, Medium Availability) cloud profiles, and FedRAMP-equivalent controls. These frameworks were not designed with AI-specific risks in mind — they address infrastructure, access control, and data handling, but do not provide guidance on:

- Model risk management and drift monitoring
- AI training data governance and lineage
- Explainability and auditability of model outputs
- Third-party AI supplier accountability
- Ongoing human oversight and override mechanisms

ISO 42001 addresses this gap directly.

2. ISO/IEC 42001:2023 — The AI Management System Standard

2.1 Overview

ISO/IEC 42001:2023 is a management system standard published by the International Organization for Standardization in December 2023. Modelled after the ISO 9001 (quality) and ISO 27001 (information security) structure, it provides a systematic, risk-based approach to managing AI throughout its lifecycle — from design and development through deployment, monitoring, and decommissioning.

An AI Management System (AIMS) built on ISO 42001 enables organizations to:

- Establish AI governance policies and accountability structures
- Identify and assess AI-specific risks and impacts
- Implement controls across the AI supply chain

Confidential - (c) 2026 SAI Cloud Secure. All rights reserved.

- Demonstrate responsible AI practices to regulators, auditors, and the public
- Achieve third-party certification against an internationally recognized standard

2.2 Structure of the Standard

ISO 42001 follows the Annex SL High-Level Structure (HLS), making it compatible with and mappable to ISO 27001 and other management system standards already deployed in government environments. The standard comprises ten clauses:

Clause	Title	Key Requirements
4	Context of the Organization	Understand internal/external context, interested parties, and AI system scope
5	Leadership	Executive commitment, AI policy, roles and responsibilities
6	Planning	Risk and opportunity identification, AI impact assessment, objectives
7	Support	Resources, competence, awareness, communication, documented information
8	Operation	AI system lifecycle controls, supply chain management, incident response
9	Performance Evaluation	Monitoring, measurement, internal audit, management review
10	Improvement	Nonconformity, corrective action, continual improvement

2.3 AI-Specific Annexes

ISO 42001 is supplemented by normative and informative annexes that provide AI-specific guidance not found in other ISO management system standards:

- Annex A — AI Controls Catalogue: 38 controls across 9 domains including AI policy, internal organization, resources, impact assessment, and AI lifecycle management.
- Annex B — Implementation Guidance: Practical guidance for applying Annex A controls across different organizational contexts.
- Annex C — AI-Specific Objectives and Risk Sources: A structured catalogue of AI risk sources including data quality, model behaviour, and third-party dependencies.
- Annex D — Interoperability with Other ISO Standards: Mapping tables to ISO 27001, ISO 9001, and ISO 31000 for integrated management system deployment.

Confidential - (c) 2026 SAI Cloud Secure. All rights reserved.

3. Mapping ISO 42001 to Canadian Government Frameworks

For Canadian federal and provincial departments, the value of ISO 42001 is amplified when it is integrated with existing compliance frameworks rather than deployed in isolation. The following mapping illustrates how ISO 42001 complements and extends existing government security and governance requirements.

Government Framework	Coverage	ISO 42001 Extension
ITSG-33 / PBMM	IT security controls, cloud profiles, Protected B data handling	AI-specific threat modelling, model risk controls, supply chain AI governance
TBS DADM	Automated decision impact assessments, transparency requirements	Lifecycle management, explainability controls, audit trail requirements
Privacy Act / Bill C-27	Personal information handling, consent, data minimization	Training data governance, AI-generated data classification, consent tracking
AIDA (Proposed)	High-impact AI systems regulation, risk classification	Risk assessment methodology, impact assessment, accountability framework
ISO 27001	Information security management system	Direct integration via Annex D; complementary control sets

Integration Insight

Organizations already certified to ISO 27001 can leverage their existing ISMS infrastructure — policies, audit programs, risk registers, and management review cycles — to fast-track ISO 42001 implementation. SAI Cloud Secure estimates a 40-60% reduction in implementation effort for ISO 27001-certified government clients pursuing ISO 42001 certification.

4. Implementing ISO 42001 in a Government Context — A Practical Roadmap

Achieving ISO 42001 compliance in a public sector environment requires a phased approach that accounts for the complexity of government IT environments, procurement cycles, and organizational change management. SAI Cloud Secure recommends the following four-phase roadmap:

Phase 1 — AI Inventory and Impact Assessment (Weeks 1–6)

- Conduct a comprehensive inventory of all AI and automated decision systems in use or under development
- Classify each system using the TBS DADM impact level criteria (I through IV)
- Identify AI systems processing Protected B or higher data classifications
- Document AI supply chain dependencies — third-party models, APIs, training data sources
- Perform ISO 42001 gap assessment against current governance practices

Phase 2 — AIMS Foundation (Weeks 7–16)

- Establish AI governance policy aligned with departmental security policy and DADM
- Define AI roles and responsibilities (AI System Owner, AI Ethics Lead, AI Security Officer)
- Develop AI risk register incorporating ISO 42001 Annex C risk source catalogue
- Implement AI supplier assessment process covering model provenance, training data disclosure, and security testing
- Create AI incident response playbooks covering model failure, adversarial attacks, and data poisoning scenarios

Phase 3 — Control Implementation and Integration (Weeks 17–28)

- Deploy AI lifecycle controls across development, testing, staging, and production environments
- Integrate AI monitoring and drift detection into existing SOC and SIEM infrastructure
- Implement explainability and audit logging mechanisms meeting DADM transparency requirements
- Conduct AI-specific threat modelling exercises incorporating MITRE ATLAS adversarial AI framework
- Integrate ISO 42001 controls with existing ISO 27001 ISMS where applicable

Phase 4 — Audit Readiness and Certification (Weeks 29–40)

- Conduct internal ISO 42001 audit against all ten clauses and Annex A controls
- Perform management review with documented AI governance KPIs and corrective actions
- Engage accredited certification body for Stage 1 (documentation review) and Stage 2 (site audit)
- Maintain continuous monitoring program with quarterly AI risk reviews and annual surveillance audits

Confidential - (c) 2026 SAI Cloud Secure. All rights reserved.

5. AI-Specific Threats in the Public Sector — What ITSG-33 Doesn't Cover

Traditional IT security frameworks like ITSG-33 are built around protecting data at rest and in transit, controlling access, and ensuring system availability. AI systems introduce a new category of threats that require distinct countermeasures:

AI Threat Vector	Description	Government Impact	ISO 42001 Control
Prompt Injection	Malicious inputs manipulate AI model outputs or instructions	Corrupted citizen-facing decisions, information leakage	Clause 8 — Input validation controls, adversarial testing
Data Poisoning	Training data is deliberately contaminated to degrade or bias model outputs	Biased automated decisions, fraud detection bypass	Annex A.6 — Training data governance and lineage
Model Inversion	Adversaries reconstruct training data from model outputs, exposing PII	Protected B data leakage from AI API endpoints	Annex A.8 — Output controls, data minimization in training
Model Drift	Model accuracy degrades over time as input distributions shift	Unreliable automated decisions affecting citizen services	Clause 9 — Performance monitoring and drift detection
Supply Chain Compromise	Third-party foundation model or dataset is compromised	Backdoor insertion into government AI systems	Annex A.5 — AI supplier assessment and contractual controls
Adversarial Examples	Carefully crafted inputs cause model misclassification	Bypass of fraud, security, or document classification systems	Clause 8 — Robustness testing and red team exercises

6. How SAI Cloud Secure Helps

SAI Cloud Secure is a cloud security advisory firm specializing in government compliance, AI security, and cloud architecture for Canadian federal and provincial clients. We bring hands-on experience with ITSG-33, PBMM cloud profiles, Protected B workloads, and emerging AI governance requirements to help public sector organizations navigate the intersection of security, compliance, and innovation.

6.1 Our ISO 42001 Advisory Services

- AI Inventory and DADM Impact Assessment — end-to-end cataloguing of AI systems with TBS DADM impact level classification
- ISO 42001 Gap Assessment — structured evaluation against all ten clauses and 38 Annex A controls with prioritized remediation roadmap
- AIMS Design and Implementation — policy development, risk register design, control implementation, and integration with existing ISO 27001 ISMS
- AI Threat Modelling — using MITRE ATLAS and government-specific threat intelligence to identify and mitigate AI-specific attack vectors
- Certification Readiness — internal audit program, management review facilitation, and pre-certification mock audit
- Continuous Compliance — ongoing monitoring, quarterly risk reviews, and annual surveillance audit support

6.2 Our Technology Stack

SAI Cloud Secure's advisory engagements are supported by a cloud-native toolset designed for Canadian government environments, including:

- Microsoft Azure Government Cloud — PBMM-compliant AI workload deployment and monitoring
- AWS GovCloud Canada — Protected B AI pipeline architecture and MLOps governance
- Anthropic Claude API — AI-assisted compliance analysis, policy drafting, and audit evidence synthesis
- SIEM/SOAR Integration — AI model monitoring integrated with government SOC infrastructure

6.3 Why SAI Cloud Secure

Capability	SAI Cloud Secure Advantage
Government Compliance Expertise	Deep knowledge of ITSG-33, PBMM, DADM, and Privacy Act requirements
AI Security Specialization	Dedicated AI threat modelling, adversarial testing, and supply chain assessment
ISO 42001 + ISO 27001 Integration	Integrated AIMS/ISMS approach reduces duplication and accelerates certification
Canadian Focus	Toronto-based, understanding of federal and provincial procurement and policy context
Agile Delivery	Lean advisory model designed for government contracting vehicles and rapid deployment

7. Conclusion and Recommended Next Steps

The convergence of accelerating AI adoption, emerging federal regulation under AIDA, and growing adversarial interest in government AI systems makes AI governance a near-term imperative for Canadian public sector organizations — not a future-state consideration.

ISO/IEC 42001:2023 provides the most credible, internationally recognized framework available today for establishing, implementing, and certifying an AI Management System. For organizations already operating under ITSG-33, PBMM, and TBS DADM obligations, ISO 42001 is a natural and additive extension of existing governance infrastructure — not a competing framework.

Recommended Next Steps for Government Organizations

1. Conduct an AI System Inventory — identify all automated and AI-assisted decision systems in use or development.

2. Perform a DADM Impact Assessment — classify systems and verify compliance with existing TBS directives.

3. Commission an ISO 42001 Gap Assessment — understand your current governance posture and prioritize remediation.

4. Engage Specialized Advisory Support — work with a partner who understands both AI security and the Canadian government compliance context.

5. Begin Phase 1 Implementation — start with policy foundation and AI risk register development within the current fiscal year.

SAI Cloud Secure is available for introductory consultations to discuss your organization's AI governance posture and how ISO 42001 can be integrated into your existing compliance program. Contact us at kp@saicloudsecure.com or visit saicloudsecure.com.

About SAI Cloud Secure

SAI Cloud Secure is a Canadian cloud security and AI governance advisory firm serving federal and provincial government clients. We specialize in ITSG-33, PBMM cloud profiles, ISO 27001/42001, and AI security architecture. Our mission is to help public sector organizations adopt cloud and AI technologies securely, compliantly, and with confidence.

Toronto, Ontario, Canada | kp@saicloudsecure.com | saicloudsecure.com

Confidential - (c) 2026 SAI Cloud Secure. All rights reserved.