KPMG

Government Internal Auditors Council of Canada | Conseil canadien des auditeurs internes du secteur public

Generative AI for Internal Audit in the Canadian Government: From Pilots to Proof

What federal, provincial, and municipal audit leaders can do now to capture value from using generative AI safely, measurably, and expediently.

Why internal audit's moment with AI is different

Internal audit (IA) in the public sector has always been about disciplined curiosity: finding out what's really happening inside complex not for profit institutions and helping leaders act on it. Generative AI (GenAI) is a new instrument for that mission – able to summarize vast evidence, surface anomalies, and draft crisp reports in minutes. But it also introduces risks that strike at audit's core strengths: independence, evidence quality, and reproducibility. AI has challenges with data quality and access, the need for human judgement and professional skepticism, issues with transparency and explainability, and risks of bias and accountability. Although it summarizes existing information, it does not interpret it or provide an opinion or conclusion.

There are a range of GenAI tools available to internal audit teams, including free online models and paid business versions, such as ChatGPT Enterprise, Microsoft Copilot M365 or Gemini Enterprise to name a few, each offering different levels of functionality, security, and internet access. Free tools may provide basic summarization and drafting features, while enterprise-grade solutions typically offer enhanced data protection, integration options, and controlled access to external sources. It is important to highlight the source of information used by each tool, as well as any limitations in functionality or data privacy. Clear distinctions between tool types help readers understand the practical and compliance implications for public sector audit work.

This article synthesizes the latest public-sector adoption signals and practical lessons from Government of Canada (GoC) internal audit teams and KPMG's recent Public Sector

AI Survey to offer IA leaders a pragmatic, risk-aware playbook tailored to Canadian contexts.

*"Only **13%** of surveyed public-sector organizations report having implemented AI; **32%** are experimenting or piloting."*

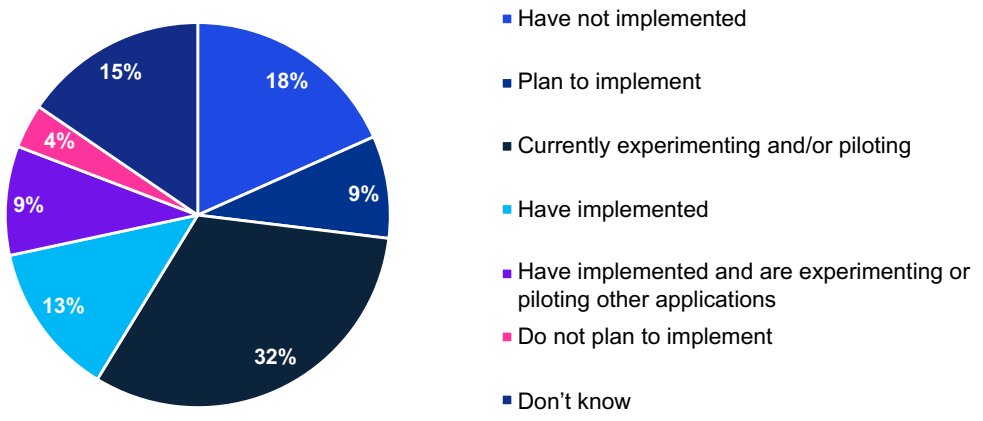
KPMG's 2025 Public Sector AI Survey.

What does KPMG's 2025 Public Sector AI Survey data tell internal audit leaders?

Across Canada's public sector, AI adoption is cautious and uneven. In a recent KPMG survey of 349 respondents across all levels of government in Canada, over half (54%) report piloted or implemented solutions. At the individual level, 52% have not used AI at work, while 24% use public tools and 16% use an employer-provided platform. Usage, when it happens, concentrates on summarizing information (64%) and drafting/editing documents (60%) – clear reproducible efficiency plays.

Figure 1 – Where organizations are with AI today

Pilots dominate; only a small minority have scaled implementation.

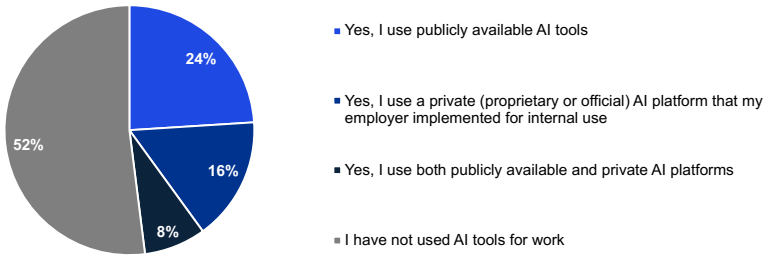


Why this matters for IA:

The dominance of pilots and experiments, with only a small minority of organizations having fully implemented AI, signals that most public sector entities are still in the "test and learn" phase. For IA, this means there is a unique opportunity to shape governance, risk management, and control frameworks before AI becomes deeply embedded. Auditors can add value by reviewing pilot project controls, advising on risk mitigation strategies, and helping organizations transition from experimentation to sustainable, scalable AI adoption. Public sector entities face some barriers in the implementation of AI including privacy, security, longer procurement timelines and quality of data to name a few.

Figure 2 – Individual use of AI at work

Over half of respondents haven't used AI at work which creates an adoption gap that IA leaders should consider in training plans.



Why this matters for IA:

With over half of respondents not using AI tools at work, there is a clear adoption gap that could lead to uneven process efficiency and inconsistent risk exposure across departments. IA should assess the root causes, such as lack of access, unclear policies, or insufficient training, and recommend targeted interventions. Audit leaders should ensure that AI adoption aligns with organizational objectives and risk appetite, while monitoring for shadow AI[1] and unauthorized tools' usage.

[1] Shadow AI refers to the use of artificial intelligence tools by employees without official approval or oversight, often bypassing organizational security and compliance controls. This can expose sensitive data and create risks that are difficult to monitor or manage within standard governance frameworks.

Figure 3 – Top barriers to AI adoption

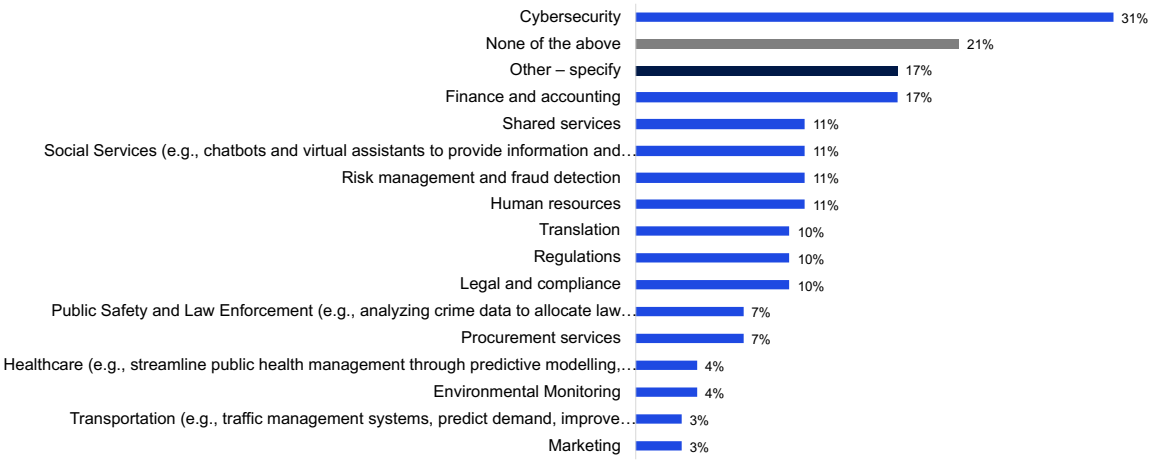
These are the domains where IA provides assurance that can be used to prioritize AI-related audit procedures.

Why this matters for IA:

The leading barriers are privacy concerns, security/compliance, and misinformation which are all domains where IA provides assurance and oversight. Lack of trust in data is also important, especially in the light of respondents indicating a lack of data policies and/or training. Auditors should prioritize reviews of AI-related privacy controls, cybersecurity measures, and data governance frameworks. By proactively addressing these barriers, IA can help organizations build trust in AI systems, reduce regulatory risk, and ensure ethical use of emerging technologies.

Figure 4 – Where do organizations plan to invest in AI (next 3 years)?

Investment aligns with IA's core audit universe (cyber, finance, risk) which acts as a cue for targeted AI assurance work.

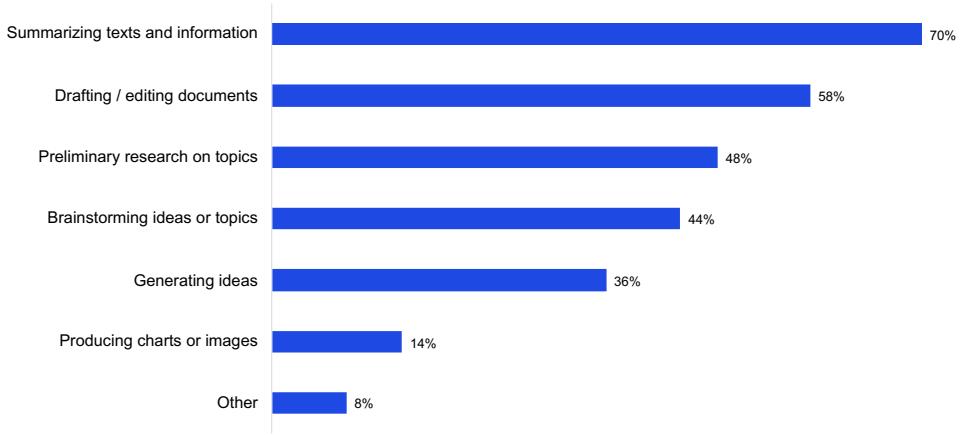


Why this matters for IA:

Planned investments in cybersecurity, finance, risk management, and shared services align closely with the core audit universe. IA should anticipate increased demand for assurance over AI-enabled processes in these areas, including controls over financial data, fraud detection, and cyber threat management. Auditors can support management by assessing the effectiveness of AI investments, identifying gaps in control environments, and recommending improvements to maximize value and minimize risk, while at the same time also utilizing GenAI tools in their audit work.

Figure 5 – What do people actually use AI for today?

AI can meet IA users where they start with summarization and drafting, then extend into structured audit artifacts.

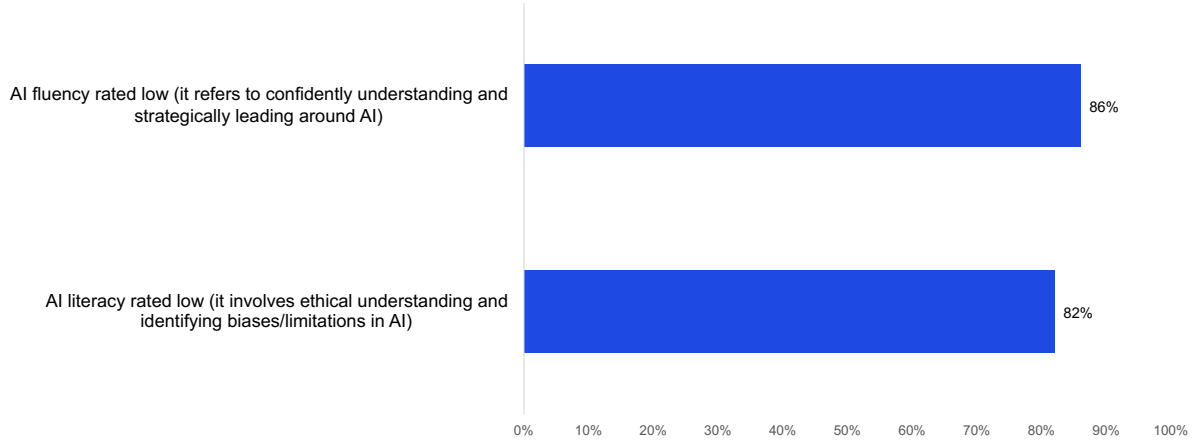


Why this matters for IA:

The most common uses of AI include summarizing information, drafting/editing documents, and conducting preliminary research which are efficiency-driven and touch core business processes. IA should assess whether these uses are governed by clear policies, whether outputs are reliable and traceable, and whether staff are trained to recognize AI limitations. Auditors can also help organizations leverage AI for enhanced reporting and analytics, while ensuring that automation does not compromise data integrity or audit trail completeness.

Figure 6 – Workforce readiness

The vast majority of respondents see both AI literacy and fluency as low among public sector employees, highlighting a critical need for targeted training and oversight.

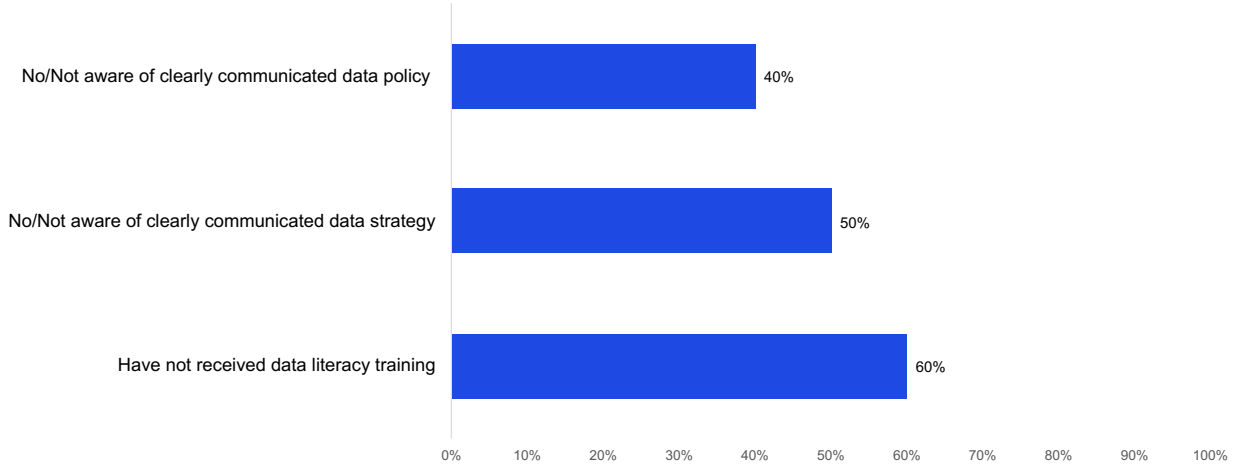


Why this matters for IA:

Low literacy and fluency heighten risks of misuse, weak oversight, and uneven adoption. IA should plan readiness baselines, prioritize training controls (e.g., role-appropriate curricula with ethical guardrails), and incorporate bias/limitations reviews into audits of AI-enabled processes.

Figure 7 – Data foundations have significant gaps

Large proportions of respondents lack or are unaware of key data policies, strategies, and training, signaling major risks for AI governance and audit assurance.



Why this matters for IA:

Weak data foundations amplify model risk (poor data lineage, consent, security) and drive inconsistent outcomes. IA should assess the presence and communication of data policies/strategy, test role-based awareness, and verify training coverage and effectiveness, making these staples of the annual risk universe.

Lessons from GoC evaluation teams that IA can leverage

Evaluation units across the GoC are already testing GenAI in real workflows. Notably, **Shared Services Canada (SSC)** used **CANChat** (an enterprise version of ChatGPT) throughout an evaluation lifecycle (planning, data collection, analysis, and reporting), while earlier SSC projects used GenAI in reporting/publication phases. Evaluators report value in **drafting interview guides**, **summarizing transcripts**, **extracting themes/codes**, and **improving clarity** of reports which are activities with strong parallels in IA.

Two operational insights stand out from these exercises:

- **Prompt discipline matters.** Small wording changes cut hallucinations. For example, instructing models to "prioritize accuracy and completeness over conciseness" can materially improve outputs.
- **Time savings aren't automatic at the engagement level.** GenAI may speed tasks, but end-to-end evaluation timelines didn't always shrink; the human value shifts to methodological rigor, contextual judgement, and defensibility which is a lesson IA should internalize. This observation aligns with the Massachusetts Institute of Technology's (MIT) recent report on the State of AI in Business, which found that 95% of GenAI pilot deployments delivered little to no measurable impact on the enterprise's profitability; however, the core issue pointed to a "learning gap" for organizations leading to flawed enterprise integration.

The evaluation community also flags **security, privacy, and ethical boundaries** which are highly relevant to audit: Government of Canada approved tools for GenAI are typically **unclassified-only** until **Authority-to-Operate (ATO)** is achieved; many evaluation contexts require **Protected B** environments which are real constraints that IA will face when audits touch sensitive personnel or financial data. Privacy practice prohibits including personal identifiers in prompts and warns against the **"mosaic effect"** (reidentification by combining innocuous data points).

International and domestic guidance converges on principles (human-centered, accountable, secure, transparent) and GoC's own **FASTER** framing (Fair, Accountable, Secure, Transparent, Educated, Relevant). IA can adapt these as **policy anchors** for tooling, workpaper standards, and training plans.

Where IA should start: high-value, low-risk use cases

1. **Planning & scoping**
    - Rapid **environmental scans** and **document summarization** (Treasury Board submissions, policies, prior audits) to formulate lines of inquiry while always storing outputs in your audit file.
    - **Preliminary control catalogues** and **risk/control matrices** seeded from public **frameworks, then tailored by auditors.**
2. **Fieldwork & evidence handling**
    - **Interview guide drafting** and **bilingual** question sets; **transcription** and **thematic coding** of notes to support consistent evidence synthesis.
    - Assisted **policy-to-control mapping** and **procedure gap scans** across large document sets where auditors validate and trace to source.
3. **Reporting & follow-up**
    - **First-draft finding write-ups/preliminary findings** (condition, criteria, cause, effect, recommendation) in a standard template; models help with clarity/consistency while auditors retain judgement.
    - **Action-plan analysis** for feasibility/clarity checks; lessons-learned summaries across audits.

Guardrails for all three: use approved, Canadian-hosted tools for unclassified data; never paste sensitive content into public models; and keep a prompt and output log in the audit file for reproducibility.

The IA control framework for GenAI (fit for Canadian governments)

To preserve **audit quality while unlocking speed embed GenAI within a lightweight but explicit control framework:**

1. **Purpose & risk classification.** Tag every GenAI task as **advisory/drafting**, **analysis-assisted**, or **evidence-generating**. Prohibit GenAI for **final judgements** or **conclusive evidence** without human corroboration.

2. **Data governance.** Define which **data classes** can be processed (unclassified vs Protected B), approved tools, and retention locations. Include checks against the **mosaic effect**.
3. **Methodological integrity.** Require **prompt repositories**, versioning, and **chain-of-thought redaction** in final deliverables while preserving working artifacts for Quality Assurance. Use the evaluation community's quality criteria (e.g., **effectiveness, equity, trust, understandability**) as acceptance gates for GenAI-assisted outputs.
4. **Security & ATO alignment.** Only tools with departmental **Authority-to-Operate**; default to **unclassified** until protected environments are available. Involve CIO/CISO early for audits touching personal or financial data.
5. **Human oversight.** Define **review tiers**: senior auditor attestation that GenAI content was fact-checked; sign-off that sources are retained and traceable.
6. **Measurement.** Track **cycle-time**, **pages reviewed per hour**, **draft-to-final edit ratios**, and **rework rates** on GenAI-assisted tasks to demonstrate Return on Investment (ROI) especially valuable where enterprise-level measurement is immature (**47%** "don't know" how ROI is measured).

Building capability: people, tools, and operating model

Upskilling focus. Canadian evaluators highlight five competency clusters that translate well to IA: digital literacy, data analysis, basic coding, data ethics/privacy, and collaboration (bridging technical and audit domains). Prioritize practical, scenario-based training – e.g., "Generate a first-draft finding and test it against our evidence standard."

Prompt library as shared infrastructure. Establish a **curated library** of prompts for audit tasks (planning memos, interview guides, testing scripts, report sections). Benefits: faster onboarding, more consistent outputs, fewer errors, and institutional memory of "what works" in your environment.

Tooling reality. Expect lag between public model capability and what's approved internally; account for license, infrastructure, and training costs noting that AI is not cost-free. Build business cases that link efficiency gains (49% target) to audit plan coverage and risk reduction.

Change management. Treat GenAI adoption as a change program: visible leadership sponsorship, clear guidelines on what's allowed, and transparent acknowledgement of AI anxiety – many staff worry about quality, fairness, or job impacts.

A pragmatic 100-day roadmap for IA

Days 0–30: "Secure the scaffolding"

- Approve a GenAI in IA policy (purpose, allowed tasks, data classes, review levels).
- Stand up an approved tool for unclassified use; create an audit file template with a GenAI log section.
- Launch micro-training (1–2 hours) on prompt hygiene and evidence standards. Share practices, lessons learned, and create an environment where experimentation is encouraged in a safe manner.

Days 31–60: "Prove the pattern"

- Run two micro-pilots: one on planning (rapid document triage) and one on report drafting (first-draft findings).
- Measure cycle-time and quality deltas vs. baseline; capture lessons in the prompt library.

Days 61–100: "Scale with guardrails"

- Expand to interview support (guide drafting, transcription, thematic coding).
- Introduce peer review rubrics mapped to quality criteria (effectiveness, trust, understandability).
- Present an ROI brief to the Audit Committee (efficiency metrics, risk posture, next-wave use cases in cyber/finance).

Conclusion: Make it real, make it safe, make it measured

Canadian institutions are leaning in carefully. Adoption is broadening from experimentation, focused on efficiency gains, even as privacy, security, and trust remain front-of-mind. IA can lead by example: pick auditable use cases, enforce clear guardrails, and measure improvements. If IA can demonstrate faster scoping, clearer findings, and

better knowledge transfer without compromising independence or evidence quality then GenAI becomes not just another tool, but a force multiplier for value.