

Performance Audit

Use of Artificial Intelligence in the Ontario Government

// Independent Auditor's Report

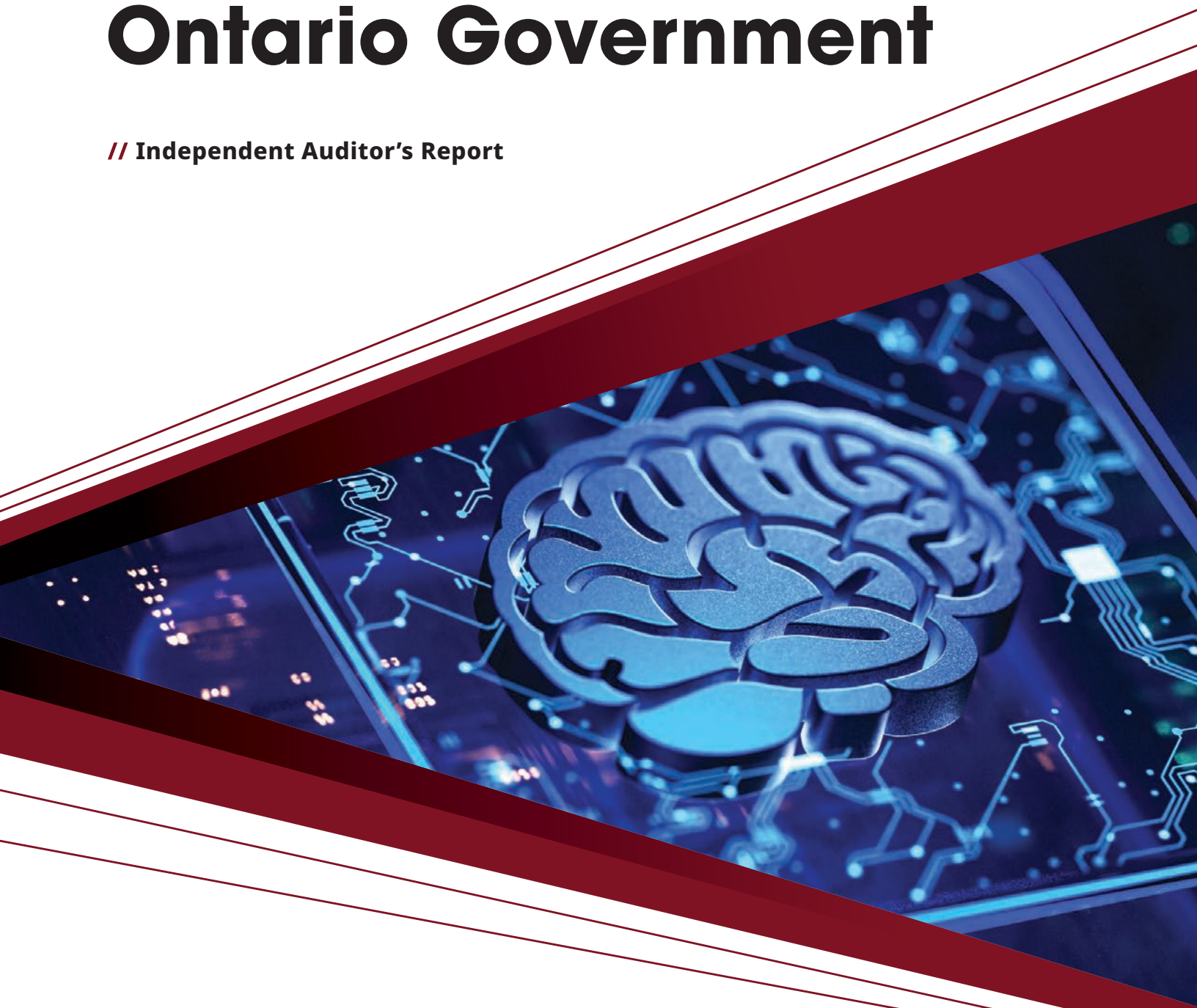


Table of Contents

1.0 Audit at a Glance	1
// Why We Did This Audit	1
// What We Found	1
// Our Conclusion	4
2.0 Background	5
2.1 Artificial Intelligence	5
2.2 AI Legislation in Ontario	5
2.3 The OPS's AI Strategy	7
2.4 Roles and Responsibilities for AI Implementation	7
2.5 The AI Framework and AI Directive	9
2.6 Key AI Systems in Use by the OPS	9
2.7 Managing Risks of AI Systems	10
3.0 Audit Objective and Scope	11
4.0 What We Found	12
4.1 Responsible Use of AI Systems	12
4.2 Risks of AI Systems	17
4.3 Safe Use of AI Scribe Systems	20
4.4 The OPS's AI Strategy and Framework	29
Recommendations and Auditee Responses	33
Audit Criteria	39
Audit Approach	40
Audit Opinion	41
Acronyms	42
Glossary	43
Appendix 1: Key AI Systems in the OPS, September 2025	45
Appendix 2: Criteria Used to Evaluate Bidders in the RFB Process for AI Scribe Systems	46



1.0 Audit at a Glance

// Why We Did This Audit

- Artificial intelligence (AI) is a rapidly evolving technology that brings both opportunities and challenges. Its use raises important concerns around transparency, privacy, security and the risk of potentially biased outcomes. AI can also impact people’s lives in positive ways by boosting productivity and innovation in various industries and in daily life.
- In November 2024, the Ontario Public Service (OPS) introduced its AI Strategy. The strategy sets out the approach for using AI to deliver services that the government provides to Ontarians. The Ministry of Public and Business Service Delivery and Procurement (Ministry) is leading the AI adoption efforts on behalf of the OPS. Our Office performed this audit to help ensure that the OPS has a strong foundation of AI principles, safeguards and controls in place.

// What We Found

OPS Staff Accessed Unsafe and Unsecured AI Websites, Creating Risks of Potential Unauthorized Data Exposure

- The Ministry had not blocked OPS staff’s access to numerous unsafe and unsecured AI websites on their OPS-provided devices.
- The Ministry had not implemented security controls to prevent OPS staff from inadvertently uploading Ontarians’ personal information or sensitive corporate data onto these AI websites.
- Of the 400 AI websites that OPS staff accessed between April 2025 and August 2025, 244, or about 60%, were deemed unsafe and unsecured according to the security score given by Microsoft’s Defender cybersecurity tool.

- As of August 2025, 3% (1,800 of 55,000) of OPS staff had completed the Ministry's Responsible Use of AI training. This training is not mandatory.

» **Recommendation 1**

Approved Microsoft Copilot Chat Had a Low Rate of Use Compared to Unapproved GenAI Websites

- Microsoft Copilot Chat is the only OPS-approved generative AI (GenAI) website, as it is in a secure environment if accessed correctly. We noted that OPS staff used it far less than other popular GenAI websites. From April 2025 to August 2025, other popular GenAI websites made up 94% of OPS staff's usage, while Microsoft Copilot Chat made up 6%.
- Although the Ministry monitored the usage of Microsoft Copilot Chat as one of its key performance indicators, it had not established targets, did not report metrics to key stakeholders, did not analyze low adoption rates and did not make efforts to ensure increased usage.

» **Recommendation 2**

OPS Staff Used Authorized GenAI on Non-default Browsers

- OPS staff are automatically signed into Microsoft Edge when they log into the OPS system, and Microsoft Copilot Chat's Enterprise Data Protection feature is then automatically enabled. This feature protects against unauthorized data disclosure. However, the feature can be bypassed if staff use Google Chrome or Mozilla Firefox without logging into their OPS account.
- When staff use Microsoft Copilot Chat on these non-default browsers, there is a risk that the data could be used by the GenAI websites to train their large language model software.

» **Recommendation 3**

More Testing Required to Evaluate the Document Verification Service AI System for Bias Risks

- The Document Verification Service (DVS) is the first external AI system to be launched by the OPS that allows Ontarians to register for online government services. Ontarians will be required to verify their identity online, using facial recognition technology, to access these services.

- We found that the Ministry did not address gaps in the DVS vendor-provided test reports. The AI Directive requires that any AI systems be trained and tested using representative data. We identified that the sample used in the testing report was too small and was not representative of the diverse demographics of Ontario's population. As a result, demographic groups may experience higher rejection rates or delays when accessing government services online.

» **Recommendation 4**

AI Scribe Systems Were Not Evaluated Adequately

- The AI Scribe request for bids (RFB) process was designed by Supply Ontario, in consultation with OntarioMD, Ontario Health and the Ministry of Health. During the AI Scribe procurement process, evaluators noted inaccuracies in the medical notes generated by most of the approved vendors' AI Scribe systems. These inaccuracies included incorrect information, AI hallucinations and incomplete information.
- We found that 11 of the 20 approved vendors did not submit any third-party audit reports, System and Organization Controls reports or International Organization for Standardization 27001 certification, and five vendors did not submit threat risk assessments and privacy impact assessments as required by the RFB process. These vendors were still approved.
- A comprehensive evaluation of vendors was not conducted to ensure that their AI Scribe systems mitigated the risk of bias.

» **Recommendations 5 to 9**

The OPS's AI Strategy Lacks Many Key Components

- The OPS's AI Strategy lacks several key components when benchmarked against AI strategy documents for public-sector organizations in Canadian and international jurisdictions, as well as selected best practices strategy documents from private sector entities. For example:
 - The Ministry did not identify specific actionable items.
 - The Ministry did not have a measured and planned approach to identify and prioritize the use of AI within various sectors or program areas in the OPS.
 - The Ministry had not explicitly identified any prohibited AI practices or areas where the use of AI may pose unacceptable risks to the public and should be banned.

» **Recommendation 10**

// Our Conclusion

Our audit found that the Ministry, on behalf of the OPS, did not have consistently effective processes and procedures in place to:

- » develop and communicate a comprehensive strategy and framework for the OPS-wide adoption of AI, supported by a governance structure to approve and monitor its consistent and responsible use;
- » completely identify, select and implement appropriate and secure AI tools and technologies; and
- » identify its workforce requirements to manage, deploy and use AI tools and technologies.

Specifically, OPS staff can access unsafe and unsecured AI websites. The Ministry had not implemented security controls to prevent its staff from inadvertently uploading personal or sensitive information to GenAI websites.

Supply Ontario, Ontario Health and OntarioMD had not comprehensively evaluated vendors' AI systems intended for use by health-care professionals.

We found that the AI Strategy lacked some key components when compared to other public-sector jurisdictions and best practices from the private sector. In addition, the Ministry had taken important steps to make its workforce AI-ready, but we found that additional improvements could be made, as their training course was not mandatory and was attended by 3% of OPS staff at the time of our audit.

The Ministry has agreed with all five recommendations addressed to it. Supply Ontario has agreed with four of the five recommendations addressed to it, and partially agreed with one recommendation.





2.0 Background

2.1 Artificial Intelligence

Artificial Intelligence (AI) is the ability of information technology systems to perform tasks that typically require human intelligence, such as analyzing large, complex data for decision-making and problem-solving based on the inputs provided, and generating audio and video content.

AI systems make decisions based on rules, algorithms or learned patterns

AI systems make decisions based on rules, algorithms or learned patterns that determine how the systems process inputs and generate outputs. AI systems can learn rules by being trained on large datasets, allowing them to learn patterns and predict what the user is expecting.

Generative AI (GenAI) is a subset of AI that generates new content based on a user's request. For example, GenAI websites can generate text, images, audio and video files, or software code, by searching the Internet to provide summarized responses to queries.

2.2 AI Legislation in Ontario

The *Enhancing Digital Security and Trust Act, 2024* (Act) came into force in January 2025. The Act enables regulation of the use of AI systems and cybersecurity by the Ontario Public Service (OPS), Crown corporations and agencies, and the broader public sector. The Act defines an AI system as “a machine-based system that, for explicit or implicit objectives, infers from the input it receives in order to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments.”

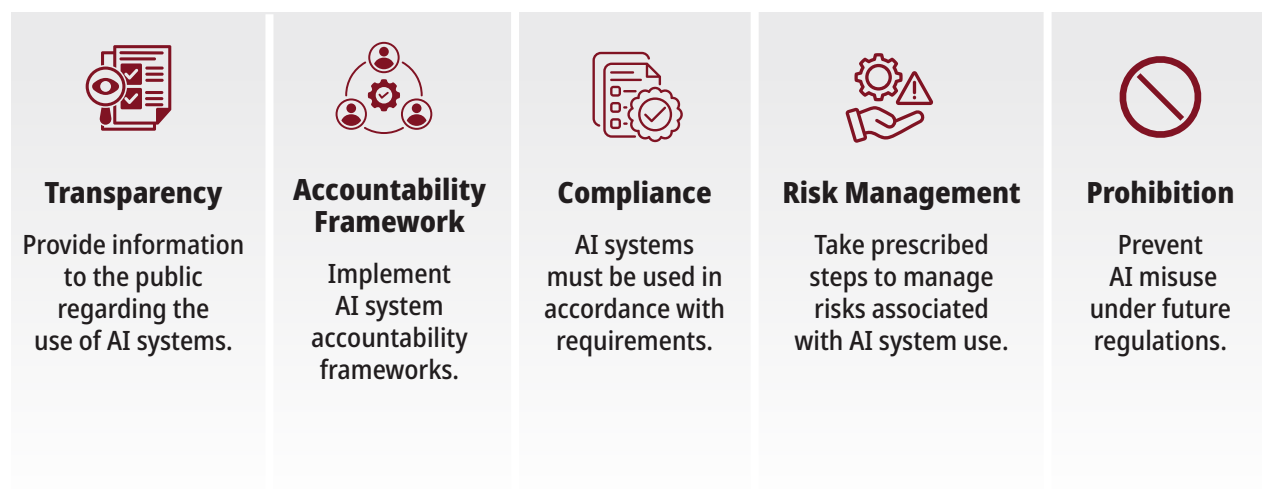


The Act grants the Minister (or designate) the authority to set technical standards and the Lieutenant Governor in Council the authority to set requirements for public-sector entities, as may be prescribed, to publish information about their use of AI in public services, establish oversight over AI and develop a responsible use framework. See the requirements of the Act in **Figure 1**.

Figure 1: Key Areas of Regulation-Making Authority, the *Enhancing Digital Security and Trust Act, 2024*

Source of data: *Enhancing Digital Security and Trust Act, 2024*

Key Areas of Regulation-Making Authority of the Act



2.3 The OPS's AI Strategy

The Ministry's mandate includes delivering on the OPS AI Framework. To support the fulfilment of this mandate, the Ministry launched an AI Strategy, "Industrialize Artificial Intelligence Across Ontario," in November 2024. It identified four priority areas, as shown in **Figure 2**.

Figure 2: Four Priority Areas of the OPS's AI Strategy

Source of data: Ministry of Public and Business Service Delivery and Procurement



Making life better in Ontario with better services

Use AI to improve service delivery and user experiences (for the OPS and the public).



Increasing AI talent and literacy across the OPS

Develop and attract top talent while upskilling the OPS.



Driving an efficient, connected OPS AI technology at scale

Reduce AI technology duplication, promoting reusable, connected technologies and prioritizing key AI investments across the OPS.



Committing to OPS-wide AI governance

Promote the responsible use of AI through AI policies, guidelines and directives.

2.4 Roles and Responsibilities for AI Implementation

The Ministry's AI mandate is executed by its Enterprise Digital and Technology Strategy Division, which is led by the Chief AI and Technology Officer. This division is responsible for the strategic planning, execution and monitoring of the OPS's AI Strategy.

The Ministry is responsible for establishing the foundational structure for AI governance. It leads the following activities:

- » defining policies, directives and guidelines;
- » developing the necessary technology and security infrastructure; and
- » supporting procurement of AI systems.

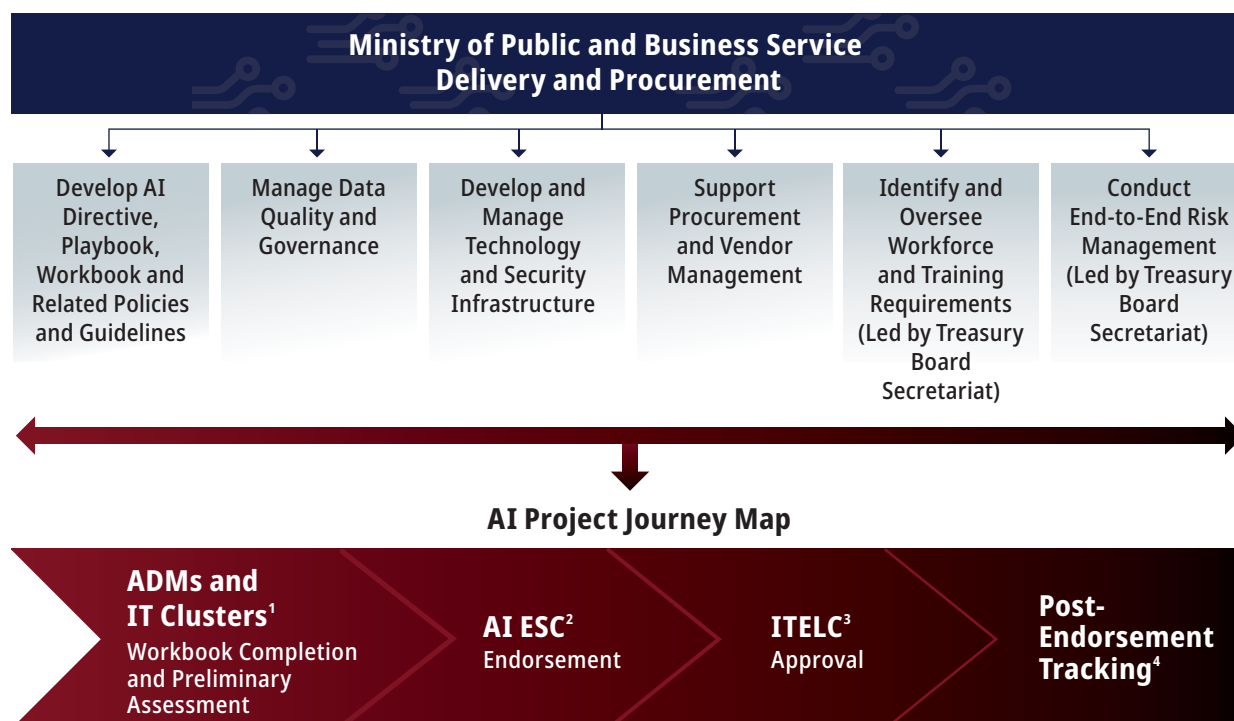
Under the governance set by the Ministry, the Treasury Board Secretariat leads the following activities:

- » identifying workforce and training needs related to AI use and deployment; and
- » supporting the assessment of risks at the OPS level, as well as for the specific AI systems in use.

The journey of an AI project from initiation to approval is encompassed within these foundational activities. Refer to **Figure 3** for a depiction of the OPS's AI governance structure and the AI project life cycle from initiation to approval.

Figure 3: The OPS's AI Governance Structure and AI Project Journey Map

Source of data: Ministry of Public and Business Service Delivery and Procurement



1. Associate deputy ministers for each ministry across the OPS, in partnership with their respective chief information officer, identify ministry-specific needs for AI use, state expected outcomes from AI solutions, conduct risk assessments and collaborate in drafting the AI Workbook in compliance with the AI Directive.
2. The AI Executive Steering Committee (AI ESC), established in December 2024, aligns the proposed AI solution to AI strategy, evaluates identified risks, assesses enterprise scalability, reviews identified AI metrics/key performance indicators, and provides leadership and support to the Ministry in fulfilling its AI mandate.
3. The Information Technology Executive Leadership Council (ITELC) is a decision-making body that sets the overall direction and priorities for the planning, development, and delivery of information and Information Technology in the Ontario government.
4. For the post-endorsement tracking process, the Ministry has not yet established roles and responsibilities, guidelines and timelines.



2.5 The AI Framework and AI Directive

The Ministry has developed an AI Framework, published in September 2023, that defines the OPS's approach to the use of AI. The framework includes the AI Directive that establishes requirements for OPS ministries and provincial agencies that use AI systems. The AI Directive, which is part of the AI Framework, includes six principles intended to guide the ethical, transparent and reliable use of AI:

- » AI is used to benefit the people of Ontario.
- » AI use is justified and proportionate, and AI systems used are reliable and valid.
- » AI is used in a safe, secure and protective way.
- » AI use is human rights–affirming and non-discriminatory.
- » AI use is transparent, and meaningful explanations of decisions are made available.
- » AI use is accountable and responsible.

Ontario ministries must show how their intended AI systems comply with these principles.

The AI Playbook is a guideline that is meant to support the OPS's compliance with the AI Directive and provides detailed explanations of how each requirement will be met. The AI Workbook is a tool designed to assist ministries in complying with the requirements of the AI Directive. Ministries complete and submit the AI Workbook with their proposed AI projects to support the review and endorsement of the OPS's AI use by the AI Executive Steering Committee.

2.6 Key AI Systems in Use by the OPS

Key AI systems were in differing phases of implementation across the OPS as of September 2025. These include Microsoft Copilot Chat, a GenAI system; REGi, an AI-powered tool that helps policy teams find and analyze regulatory requirements; and Document Verification Service, a system to verify identities and enable users to subsequently access government services.

As noted in **Figure 3**, the ministries select AI systems through a process that considers the benefits and risks, including due regard for economy, compared to similar systems. See **Appendix 1** for details on these AI systems.



2.7 Managing Risks of AI Systems

The Ministry's Cyber Security Division conducts assessments to ensure the security, privacy and compliance of data in AI systems. It also monitors the OPS environment, including network activity, to track websites accessed by staff.

2.7.1 Risk Assessments

According to the AI Directive, all AI systems used by the OPS must undergo a risk assessment process to identify and mitigate risks related to bias, transparency and cybersecurity weaknesses. Additionally, AI systems should be assessed for risks related to the collection and processing of personal or sensitive information.

Threat risk assessments (TRAs) are critical for identifying and remedying cybersecurity weaknesses within an IT system before it is implemented or undergoes a significant change. TRAs can help determine whether IT system weaknesses and threats have been adequately considered and addressed and help ensure that risks are appropriately rated based on the severity of the threats and the confidentiality of the underlying data within the IT system.

Privacy impact assessments (PIAs) identify and evaluate potential privacy risks associated with an organization's collection, use and disclosure of personal information, such as patient diagnoses, medical history, and audio recordings of conversations between patients and health-care professionals. PIAs assess the likelihood and impact of these risks to help inform decision-making for managing risks and implementing controls to prevent, or at least reduce, privacy breaches.

2.7.2 Monitoring of AI Systems and Websites

The Cyber Security Division is responsible for monitoring OPS-wide IT systems and networks to detect unauthorized access and malware activity. It has implemented some security tools, such as Microsoft Defender and Microsoft Copilot Chat's Enterprise Data Protection feature, to monitor the websites that are accessed by OPS staff in an effort to ensure safe use of AI in the OPS. Microsoft Defender is specifically used for tracking:

- » access to popular GenAI websites by OPS staff;
- » the file size or length of query that is being sent to AI websites, such as the length of a question or the size of a document being uploaded; and
- » the number of staff accessing AI websites and their Internet Protocol addresses, which are unique numbers assigned to each device that help identify which device is accessing a site.

OPS has entered into an agreement with Microsoft to make sure that the data entered and processed using Microsoft Copilot Chat resides in Canada and that the data cannot be accessed by other organizations.



3.0 Audit Objective and Scope

Our audit objective was to assess whether the Ministry, on behalf of the OPS, has effective processes and procedures in place to:

- » develop and communicate a comprehensive strategy and framework for the OPS-wide adoption of AI, supported by a governance structure to approve and monitor its consistent and responsible use;
- » identify, select and implement appropriate and secure AI tools and technologies; and
- » identify its workforce requirements to manage, deploy and use AI tools and technologies.

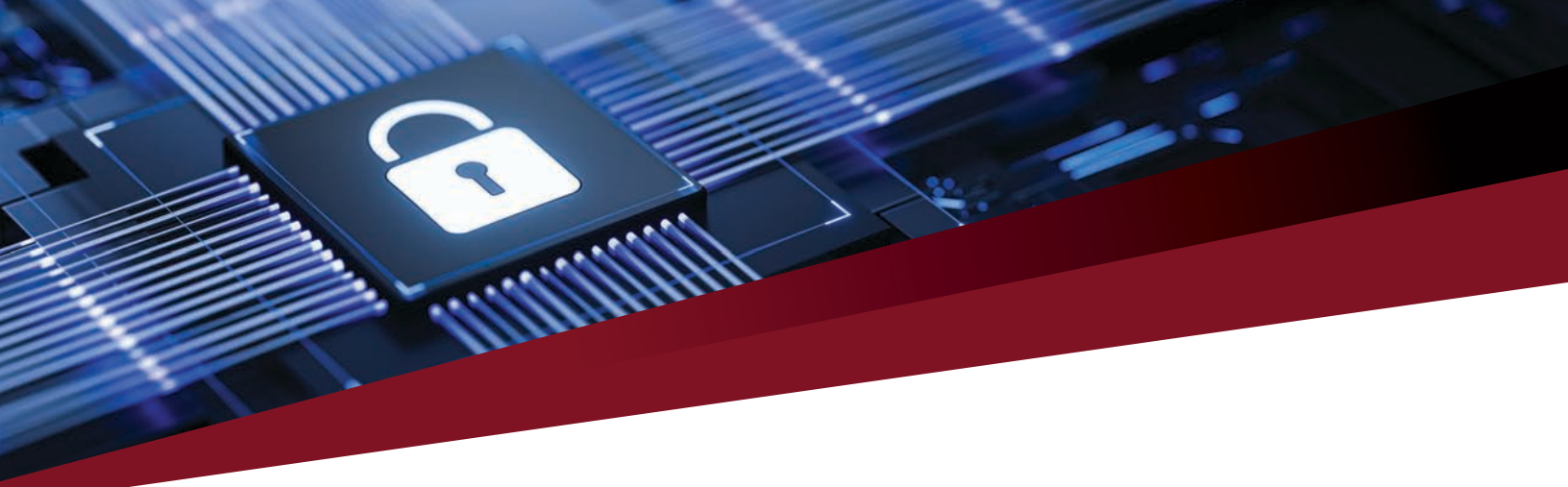
The scope included a review of the Ministry's processes related to the implementation of controls to address key risks related to cybersecurity, data residency, data privacy and AI bias while selecting vendor-developed AI systems. We also examined controls to restrict and monitor the use of unapproved or vulnerable AI systems and websites by OPS staff.

Our scope included benchmarking OPS's AI Strategy, governance approach and responsible use of AI systems against leading practices across Canadian and international jurisdictions, including a review of relevant frameworks, policies, risk management standards and publicly available AI-related incidents.

The audit also examined the procurement of AI Scribe systems used in the broader public sector by reviewing the requirements of the request for bids (RFBs) and related vendor submissions.

Our audit covered the period from January 2025 to November 2025.

For more details, see our **Audit Criteria**, **Audit Approach** and **Audit Opinion**.



4.0 What We Found

4.1 Responsible Use of AI Systems

4.1.1 OPS Staff Accessed Unsafe and Unsecured AI Websites, Creating Risks of Potential Unauthorized Data Exposure

We found that the Ministry had not blocked OPS staff's access to numerous unsafe and unsecured AI websites on their OPS-provided devices.

~ 60%

**of AI websites accessed
by OPS staff were deemed
unsafe or unsecured**

The Ministry had not implemented security controls to prevent OPS staff from inadvertently uploading Ontarians' personal information, such as their health cards, driver's licences and credit card information, or sensitive corporate data, such as vendor contracts and invoices, onto these AI websites. When OPS staff use publicly available GenAI websites, there is a risk

that these websites can retain and use the data or any personal or sensitive information entered by staff to train the sites' large language model (LLM) software.

The Ministry had controls in place to identify the websites accessed by OPS staff, but it did not have controls to monitor activities or actions performed by OPS staff on those websites, such as uploading personal or sensitive information.

One tool that the OPS uses is Microsoft Defender, which scans the OPS network to discover AI websites used by OPS staff. It assigns the websites a security score from zero to 10, where a low score indicates a higher risk of data exposure.

Between April 2025 and August 2025, the latest available reporting period, 12,000 OPS staff accessed approximately 400 AI-related websites. Of these websites, 244, or about 60%, were deemed unsafe or unsecured since they had a score of five or lower, and not all of these websites were work-related. We found that 15% of these websites with a score of five or lower were also not work-related and featured inappropriate content.

We also noted that one AI website was reported to have privacy and security concerns in early 2025. In February 2025, the federal and BC governments blocked access to the website on their networks and devices. The Ministry issued an advisory to OPS staff on the risks associated with the website that month, but didn't block access to it until March 2025.

Three Percent of OPS Staff Completed AI Training

The Ministry launched an AI training course, Responsible Use of AI, in January 2024. As of August 2025, 1,800 of 55,000 (3%) of OPS staff had completed it. We noted that the training is comprehensive and covers the use of Microsoft Copilot Chat, including the type of data that may and may not be uploaded onto it and the associated risks. In addition, the training also covers AI concepts such as GenAI, safe use of AI websites, and the risks of misinformation, bias and security issues. It also warns against using publicly available AI tools due to privacy and confidentiality risks. Although the training is available to all staff on the OPS's internal website, it is not mandatory.

■ Why It Matters

As the OPS begins to implement AI systems, it is essential that it has a strong foundation of AI principles, safeguards and controls in place to use AI to its full potential while protecting Ontarians' personal and sensitive information.

By allowing staff to use unapproved, unsafe and unsecured AI websites, and not preventing the inadvertent uploading of data to these sites, the Ministry has not implemented sufficient controls to prevent potential serious data misuse by third parties. Even a single incident of one employee uploading personal or sensitive data or clicking a malicious link on these websites can lead to data exposure, credential theft and system outage.

Training in the use of AI is also vital to enforce the use of approved platforms, like Microsoft Copilot Chat, and implement data loss prevention tools to help ensure that personal or sensitive information is not inadvertently uploaded to unsafe and unsecured AI websites.

Recommendation 1

We recommend that the Cyber Security Division of the Ministry:

- review and block OPS staff's access to unsafe and unsecured AI websites; and
- implement measures to ensure that all staff take the training on AI-related risks with refresher training offered as the OPS advances its AI use.

For the auditee responses, see **Recommendations and Auditee Responses**.



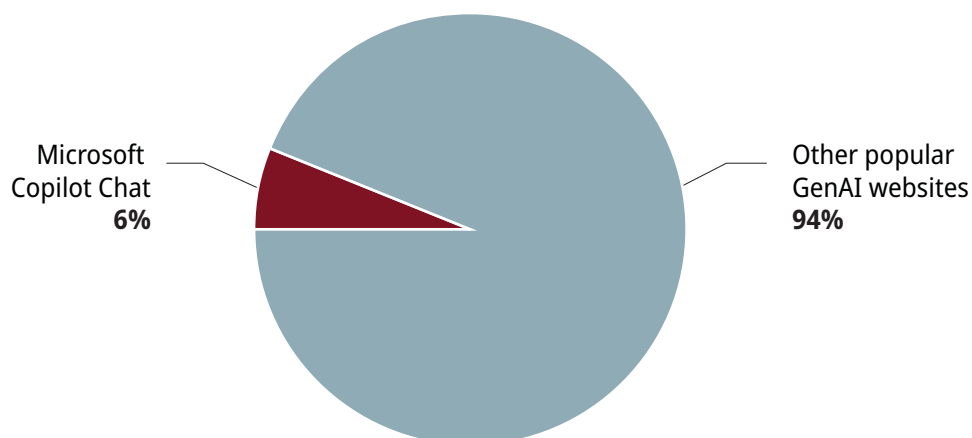
4.1.2 Approved Microsoft Copilot Chat Had a Low Rate of Use Compared to Unapproved GenAI Websites

Microsoft Copilot Chat was made available to all OPS staff in October 2024 to research, summarize information, generate content and write software code. It is the only GenAI website approved by OPS and includes the default Enterprise Data Protection feature, which ensures that any information or queries/prompts entered by OPS staff remain within the OPS's secure environment. Its rate of use by OPS staff remained significantly lower than that of other popular GenAI websites, which do not have a similar feature, making them less secure than Microsoft Copilot Chat for organizational use.

The Ministry's tracking of GenAI website usage statistics shows that, from April 22, 2025, to August 18, 2025, other popular GenAI websites made up 94% of OPS staff's usage and Microsoft Copilot Chat made up approximately 6%, as shown in **Figure 4**.

Figure 4: OPS Staff Usage of Microsoft Copilot Chat Versus Other Popular GenAI Websites, April 2025–August 2025

Source of data: Ministry of Public and Business Service Delivery and Procurement



Many private-sector organizations, including financial institutions, insurance companies and auditing firms, have restricted employee access to all GenAI websites except those approved by the organization. In contrast, OPS staff continued to be able to access other GenAI websites that serve similar purposes.

The Ministry Had Not Defined KPI Targets to Help Enforce Usage of Microsoft Copilot Chat to Reduce Risk of Data Exposure

Key performance indicators (KPIs) are established metrics used to measure an organization's effectiveness in achieving its strategic and operational goals. In the context of AI systems, KPIs provide the Ministry's management with insight into the value derived from its investments in AI.

Management can use KPIs to monitor how the AI systems are being used by staff, such as the type of tasks performed, the number of staff using AI and whether AI is enhancing the efficiency of tasks or reducing the time taken to complete them. KPIs can help promote transparency to Ontarians in how management is using AI to deliver public services efficiently and how management remains accountable for outcomes from AI systems that impact Ontarians.

KPI targets can help monitor how the AI systems are being used by staff

We noted that, although the Ministry was monitoring the usage of Microsoft Copilot Chat as one of its KPIs, it had not set any targets. While the Ministry tracked Microsoft Copilot Chat usage, the absence of defined KPIs made it difficult to determine whether usage levels demonstrated effective adoption or use of AI tools.

For an organization that aims to be "most AI-enabled government in Canada," KPI targets are critical to measure the intended objective. In addition, the Ministry did not share these usage metrics or conduct regular meetings with the Treasury Board Secretariat, or the Corporate Chief Information Officer and the Deputy Minister of the Ministry, to review why the usage of Microsoft Copilot Chat was lower than that of other popular GenAI websites.

■ Why It Matters

Microsoft Copilot Chat has safeguards and data protection that other popular GenAI websites do not. When OPS staff use GenAI websites other than Microsoft Copilot Chat, the security of any data entered as a query or prompt may be in question. To safeguard data, the Ministry needs to ensure Microsoft Copilot Chat is the only GenAI website being used by OPS staff.

Recommendation 2

We recommend that the Ministry:

- establish KPI targets to measure and track Microsoft Copilot Chat's adoption;
- take actions to increase use of Microsoft Copilot Chat to the targeted rates and usage in the OPS; and
- report these KPIs to management on a monthly basis.

For the auditee responses, see **Recommendations and Auditee Responses**.



4.1.3 OPS Staff Used Authorized GenAI on Non-default Browsers

The OPS's login settings automatically sign staff into Microsoft Edge with their OPS credentials, which in turn enables Microsoft Copilot Chat's Enterprise Data Protection feature whenever staff access the website. This feature helps protect against unauthorized data disclosure by:

- » deleting prompts and responses from the chat history after a certain time;
- » preventing Microsoft from accessing or viewing the data entered, including any personal or sensitive OPS data; and
- » ensuring that the data is not used to train Microsoft Copilot Chat's LLM software.

Non-default browsers, such as Google Chrome and Mozilla Firefox, do not always have this protection even when accessing Microsoft Copilot Chat.

When OPS staff use Microsoft Copilot Chat on non-default browsers, such as Google Chrome and Mozilla Firefox, this feature can be bypassed and there is a risk that these websites can retain and use the data or any personal or sensitive information entered by staff to train the sites' LLM software.

■ Why It Matters

If the Ministry cannot track when OPS staff use non-default browsers, any personal or sensitive information that OPS staff enter could be stored to train the site's LLM software and could be used by the browsers to improve their AI systems. This may unintentionally expose Ontarians' confidential information.

Recommendation 3

We recommend that the Cyber Security Division of the Ministry:

- block access to using Microsoft Copilot Chat on other browsers; and
- educate OPS staff through AI training about the dangers of using non-Microsoft browsers when accessing AI websites.

For the auditee responses, see **Recommendations and Auditee Responses**.



4.2 Risks of AI Systems

4.2.1 More Testing Required to Evaluate the Document Verification Service AI System for Bias Risks

The Ministry procured the Document Verification Service (DVS), which will be the first AI system to be launched externally by the OPS for Ontarians' use. DVS will allow Ontarians to register for online government services. As part of the registration process, Ontarians will be required to verify their identity online by taking a live video and performing actions such as smiling or moving their face to show that they are a real person. DVS uses an underlying facial recognition technology for facial

DVS uses facial recognition technology to verify a user

matching and to detect liveness to verify a user. Once these checks are complete, the user will be able to access government services or records. In the future, DVS is expected to enable Ontarians to verify their identities to access government services online, such as social assistance or health-related programs.

We found that the Ministry did not address gaps in the DVS vendor-provided test reports as of November 2025. The AI Directive requires that any AI system be trained and tested using representative data. However, we identified that the sample used in testing was too small and was not representative of the diverse demographics of Ontario's population.

This omission could leave the system open to generating decisions that disadvantage certain demographic groups since it uses technologies such as the facial recognition of different demographic users. As a result, certain groups may experience higher rejection rates or delays when verifying their identities online to access government services.

We reviewed the two vendor-provided testing reports, dated November 2022 and February 2023, and noted many gaps, as well as non-compliance with the requirements of the AI Directive, as shown in **Figure 5** below.

These critical limitations in the vendor-provided test reports were not challenged by the Ministry, and, at the time of our audit, there was no plan in place to conduct more representative testing or to monitor bias risks after deployment.

Figure 5: Our Office’s Evaluation of AI Directive Requirements to Assess and Mitigate Potential Bias in AI Systems

Source of data: Ministry of Public and Business Service Delivery and Procurement

AI Directive Requirement	Detailed Explanation	Our Office’s Evaluation
<p>1. Select products which are trained using relevant data (e.g., representative of the Canadian population and demographics).</p>	<p>Ministries should procure AI systems that have been trained on data representative of Ontario’s population.</p>	<ul style="list-style-type: none"> » The 2022 vendor report noted that the testing group of 214 individuals was “too small,” and not sufficiently diverse and representative of the Canadian population and demographics. » The 2023 vendor testing of the spoof or presentation attack detection engine was based on a limited sample, covering two age cohorts (18–30 years and 51–70 years).
<p>2. Select products which can transparently describe reasoning and/or confidence.</p>	<p>Ministries should procure AI systems that can clearly explain the reasons for their decisions and the factors they considered to generate outcomes, to enable users to understand, challenge or trust the outcomes generated by the AI system.</p>	<ul style="list-style-type: none"> » The Ministry procured a proprietary product that lacked transparency in how the product made decisions or the reasoning behind the decisions.
<p>3. Validate output by performing user testing and exploratory testing with diverse users.</p>	<p>Ministries should procure AI systems that are tested with a wide range of users to confirm accuracy and fairness.</p>	<ul style="list-style-type: none"> » The Ministry did not fully validate the vendor’s 2022 and 2023 testing reports, where we identified several gaps as noted in Requirement 1 above. The Ministry did not sufficiently evaluate and address these gaps.



■ Why It Matters

One of the key principles underlying the AI Framework and Directive is that “AI use is human rights–affirming and non-discriminatory.” This means that outcomes from the use of AI systems should be fair and the risk of bias should be mitigated. Without comprehensive, appropriate bias evaluation and representative testing, there is an increased risk that the AI systems may produce discriminatory or inaccurate results, inequitable access to government services, reduced trust in AI-enabled services, and non-compliance with the OPS’s established AI principles and Directive requirements.

Recommendation 4

We recommend that the Ministry validate test results provided by AI system vendors to ensure the testing was performed using a sufficient sample set representing Ontario’s demographics and that the testing meets all of the AI Directive requirements.

For the auditee responses, see **Recommendations and Auditee Responses**.



4.3 Safe Use of AI Scribe Systems

The AI Scribe program was initiated and partly funded by the Ministry of Health for use across the broader health sector by physicians, family doctors, nurse practitioners, therapists and other health professionals. Supply Ontario, a provincial procurement agency, established a vendor of record (VOR) arrangement to allow health-care providers to purchase AI Scribe systems from pre-qualified vendors.

Health-care providers can purchase AI Scribe systems from vendors pre-qualified by Supply Ontario

We reviewed the RFB procurement process conducted by Supply Ontario, in consultation with OntarioMD, Ontario Health and the Ministry of Health, and noted significant gaps related to AI Scribe systems, including the inaccurate capture of conversations, poor cybersecurity controls and lack of controls to mitigate bias risks. These gaps are detailed in **Sections 4.3.1 to 4.3.5**.

The RFB process conducted by Supply Ontario occurred in two stages. The first qualification stage required vendors to self-attest that their AI Scribe systems complied with certain requirements, such as security controls and generation of comprehensive notes. Vendors who attested that they complied with these requirements were moved to the second stage, where they were assigned scores based on key criteria, as detailed in **Appendix 2**.

It is optional for broader-public-sector entities to adopt the Supply Ontario-procured AI Scribe systems. These entities may still procure and use unapproved AI Scribe systems from vendors that are not part of the VOR arrangement.

4.3.1 Supply Ontario’s Vendor Scoring Approach Assigned Low Weightings for Critical Criteria Such as Security, Accuracy and Bias Controls

We found that some important criteria were assigned lower weightings within Supply Ontario’s RFB scoring approach. As shown in **Figure 6**, the “System security controls” criteria, such as TRAs, PIAs and System and Organization Controls (SOC) reports, accounted for 2% and 4%, respectively, of the maximum possible score; the “Accuracy of medical notes generated” criteria accounted for 4%; and bias controls within the “Data privacy/legal controls” criteria accounted for 2%. **Appendix 2** details the criteria used to evaluate bidders in the RFB process.

Figure 6: Supply Ontario RFB Stage 2 Criteria, with Maximum and Average Actual AI Scribe Vendor Scores

Source of data: Supply Ontario

Criteria	Maximum Score per RFB	Weighting or Score %	Actual Average Score of 20 Approved Vendors
Domestic presence in Ontario	159	30	136
Data privacy/legal controls	120	23	83
Bias controls	10	2	6
System security controls	63	11	45
TRAs and PIAs	10	2	3
SOC 2 Type 2	20	4	7
Contract negotiation, including vendor-requested amendments	53	10	47
Clinical format, including transcription of discussions between multiple speakers	47	9	43
Business, including system interface, vendor onboarding, training and pricing	28	5	23
Accuracy of medical notes generated	20	4	12
Total	530	100	405



No minimum passing scores were assigned to the criteria in the second stage, enabling a bidder to score zero for the “System security controls” criteria, the “Accuracy of medical notes generated” and the bias controls sub-criteria at this evaluation stage and still meet the minimum aggregate score of 371 points required to be approved as a VOR.

■ Why It Matters

Critical criteria, such as system security, privacy (including TRAs, PIAs and SOC reports), accuracy and bias controls are essential when procuring an AI system. Inadequate weightings could result in the selection of vendors whose AI tools may produce inaccurate or biased medical records or lack adequate protection to safeguard sensitive personal health information.

Recommendation 5

For all future procurements related to AI systems, we recommend that Supply Ontario:

- increase the weighting assigned to criteria related to security and privacy controls, such as TRAs, PIAs and SOC reports, bias, and accuracy; and
- assign a minimum passing score or thresholds for these criteria.

For the auditee responses, see **Recommendations and Auditee Responses**.

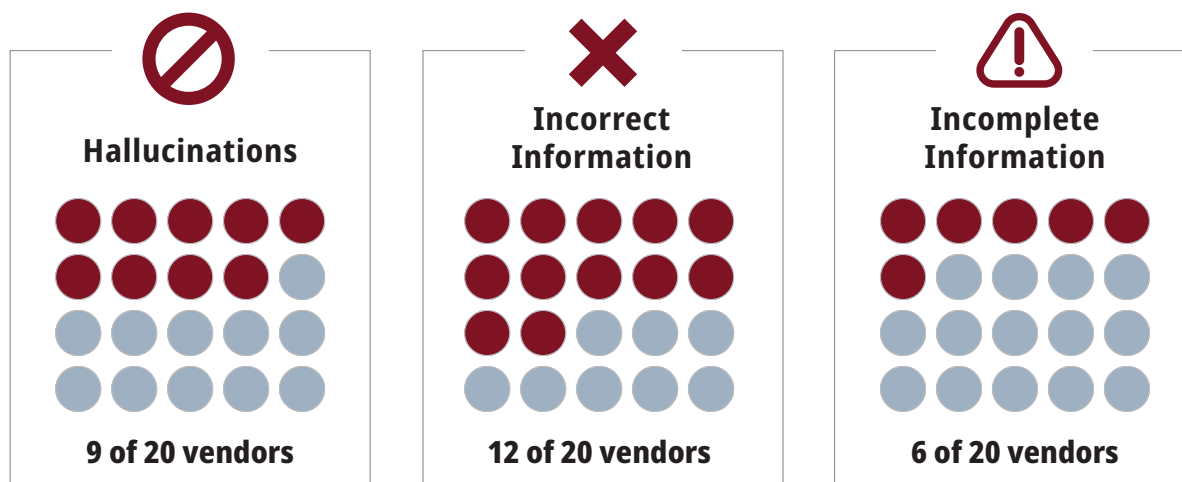
4.3.2 AI Scribe Systems Inaccurately Transcribed Details of Conversations Between Health-Care Professionals and Patients During Procurement

As part of the “Accuracy of medical notes generated” criteria in the RFB process, Supply Ontario conducted two transcription tests to assess whether the bidding vendors’ AI Scribe systems could accurately and completely record and transcribe conversations between health-care professionals and their patients and generate structured notes. Supply Ontario provided the same two simulated recordings to all vendors. The vendors’ system-generated notes were then submitted to medical professionals from OntarioMD and Ontario Health who were designated as evaluators to assess whether the structured notes accurately and completely summarized the interaction.

The evaluators’ comments highlighted inaccuracies in these system-generated notes. The inaccuracies had three themes, and all 20 of the approved vendors showed one or more of these types of inaccuracies in their system-generated notes, as shown in **Figure 7**.

Figure 7: Types of Inaccuracies Found in Notes Generated by AI Scribe for 20 Approved Vendors

Source of data: Supply Ontario



We reviewed the evaluation comments and noted the following:

Hallucinations: Nine of 20 (45%) AI Scribe systems fabricated information and made suggestions to patients’ treatment plans, such as referring the patient for therapy or ordering blood tests, even though these steps were not mentioned in the simulated recordings. Evaluators also noted hallucinations that could impact patients’ health. For example, the submitted notes included statements that there were “no masses found” or that there was presence of anxiety in the patient, although this information was not discussed in the recordings.

Incorrect information: Evaluators found that notes generated by 12 of 20 (60%) AI Scribe systems captured a different drug than what was prescribed by the doctor.

Missing or incomplete information: Notes generated by 17 of 20 (85%) AI Scribe systems missed key details about the patients’ mental health issues in at least one of the two tests, even though this was mentioned in the simulated recordings. In addition, six of 20 (30%) vendors’ system-generated notes missed the patients’ mental health issues fully or partially or were missing key details about those issues across both tests.



OntarioMD issued guidelines to doctors for the manual review of the system-generated notes to ensure they are appropriate and mitigate the risks of inaccurate outputs from AI Scribe systems. Doctors were not required to attest that they had verified the system-generated notes through a sign-off feature in the AI Scribe systems.

60%

of approved AI Scribe systems generated notes capturing a different drug than what was prescribed, during procurement evaluation

We noted that the UK has implemented regulations and guidelines to minimize inaccuracies in AI Scribe system-generated notes. National Health Services England issued specific guidance related to AI Scribe systems in April 2025, directing National Health Services organizations, such as hospitals and health-care professionals, to only use systems that have been assessed for safety and compliance through registering them with the United Kingdom’s Medicines and

Healthcare products Regulatory Agency as Class I medical devices. In addition, the National Health Services requires vendors to comply with clinical requirements to establish accuracy of the systems.

■ **Why It Matters**

Inaccuracies in medical notes generated by AI Scribe systems could potentially result in inadequate or harmful treatment plans that may potentially impact patient health outcomes. It is important that the AI Scribe systems are tested to provide assurances as to the quality of their generated notes and to minimize inaccuracies.

Recommendation 6

We recommend that Supply Ontario:

- review industry standards and guidelines from other jurisdictions related to AI Scribe systems and implement best practices in Ontario; and
- require the AI Scribe system vendors to implement IT controls within their AI Scribe systems to enforce an attestation from users of the AI Scribe to confirm their review of notes produced.

For the auditee responses, see **Recommendations and Auditee Responses**.



4.3.3 Vendors Failed to Provide Security Assessment Reports, Creating a Risk of Potential Exposure of Ontarians' Health Data

As part of the RFB process, participants were required to provide either security documents, such as internal policies and external audit reports, or a timeline for when they could provide these documents. External audit reports, such as SOC 2 Type 2 reports, HITRUST certification or International Organization for Standardization (ISO) 27001 certification, and other reports such as TRAs and PIAs, demonstrate that vendors had implemented controls to address privacy and security risks of their AI Scribe systems. ISO 27001 is the international standard for managing information security, providing a risk-based framework to protect the confidentiality, integrity and availability of an organization's data.

11 of 20

Approved vendors did not submit third-party audit reports

We noted several gaps in enforcing these requirements as part of the RFB process. Of the 20 approved vendors, 11 did not submit SOC reports, HITRUST certification or ISO 27001 certifications, and five did not submit TRAs and PIAs. All the vendors provided declarations that they maintain the necessary documentation and reports.

In addition, we noted that the evaluation process did not assess whether the SOC 2 Type 2 reports included the expected controls to address AI risks. The evaluation focused only on whether the reports were submitted on time, in accordance with the RFB requirements, and whether the reports noted any exceptions.

The evaluation process relied solely on vendors' confirmation that the data is processed and stored in Canada, without requiring supporting evidence or independent verification.

These gaps indicate that Supply Ontario, in consultation with OntarioMD and Ontario Health, primarily relied on vendors' self-attestation of compliance rather than obtaining and reviewing supporting documentation related to security and privacy requirements to safeguard Ontarians' personal health data. This approach allowed vendors to potentially overstate their compliance with security and privacy requirements.

The AI Scribe procurement process primarily relied on vendors' self-attestation of security and privacy requirements

In December 2024, a privacy breach under the *Personal Health Information Protection Act* was reported to the Office of the Information and Privacy Commissioner of Ontario regarding a hospital in Ontario that experienced a data privacy breach in September 2024. An unapproved AI transcription tool that was used by a former staff member had recorded personal patient information discussed at a meeting and automatically

distributed the transcribed notes to current and former staff. This data privacy breach occurred prior to implementation of the VOR arrangement for AI Scribe systems in April 2025.

■ Why It Matters

When Ontarians see their doctor, they need to share intimate information about their health, their bodies and their personal lives to receive proper care. Ontarians expect this extremely personal information to be kept private and confidential. Using AI to assist in providing health care must not come at the cost of compromising privacy.

Recommendation 7

We recommend that Supply Ontario, in collaboration with the Ministry:

- as part of its continuous vendor management process, obtain third-party reports, such as SOC 2 Type 2 reports and security assessment reports, from all AI Scribe system vendors annually;
- when procuring software of any type, including AI, always require a SOC 2 Type 2 report and/or security assessment to be included in the vendor response for assessment; and
- ensure the evaluator assesses whether the security assessment reports included the expected controls and review any exceptions in the reports for potential risks.

For the auditee responses, see **Recommendations and Auditee Responses**.



4.3.4 Potential Risks Related to Bias in AI Scribe Systems Were Not Comprehensively Evaluated

Supply Ontario, OntarioMD and Ontario Health did not perform a comprehensive evaluation of vendors to ensure that their AI Scribe systems mitigate the risk of generating unfair or biased outcomes.

Bias in AI Scribe systems can occur, for example, through misinterpretations of conversations between health-care professionals and patients, particularly if either party speaks with a diverse accent. This can cause the AI Scribe systems to misinterpret important health details, which could potentially lead to inaccuracies in system-generated notes.

We noted that, as part of the RFB process related to the “Data privacy/legal controls” criteria, Supply Ontario required participating vendors to describe their organizational processes relating to controls to mitigate potential bias in their AI Scribe systems. They did not require the participating vendors to provide evidence of controls, such as bias testing results.

■ Why It Matters

This limited approach to evaluating bias risks could impact the fairness, equity and reliability of the AI Scribe systems qualified by Supply Ontario and the outputs they generate. Not testing for bias could result in health-care professionals receiving inaccurate or biased system-generated notes, which may result in adverse health-related outcomes for patients.

Recommendation 8

For all future procurements related to AI systems, we recommend that Supply Ontario follow the principles of AI Directive and evaluate bias risks related to AI systems by requiring evidence of bias testing results or doing their own bias testing prior to selecting a system.

For the auditee responses, see **Recommendations and Auditee Responses**.



4.3.5 Vendors Were Not Required to Provide Live Demos of Their AI Scribe Systems

We found that Supply Ontario did not require AI Scribe vendors to demonstrate their systems live or operate them in front of the evaluators. The vendors were provided with the simulated recordings mentioned in **Section 4.3.2** and were then allowed to generate structured notes offline and subsequently submit them to Supply Ontario, Ontario Health and OntarioMD for evaluation. While vendors were required to sign an attestation confirming that the submitted notes were unedited, the absence of a live demo created a potential risk that vendors could process the recordings multiple times or alter the system-generated notes, compromising the integrity of the evaluation process.

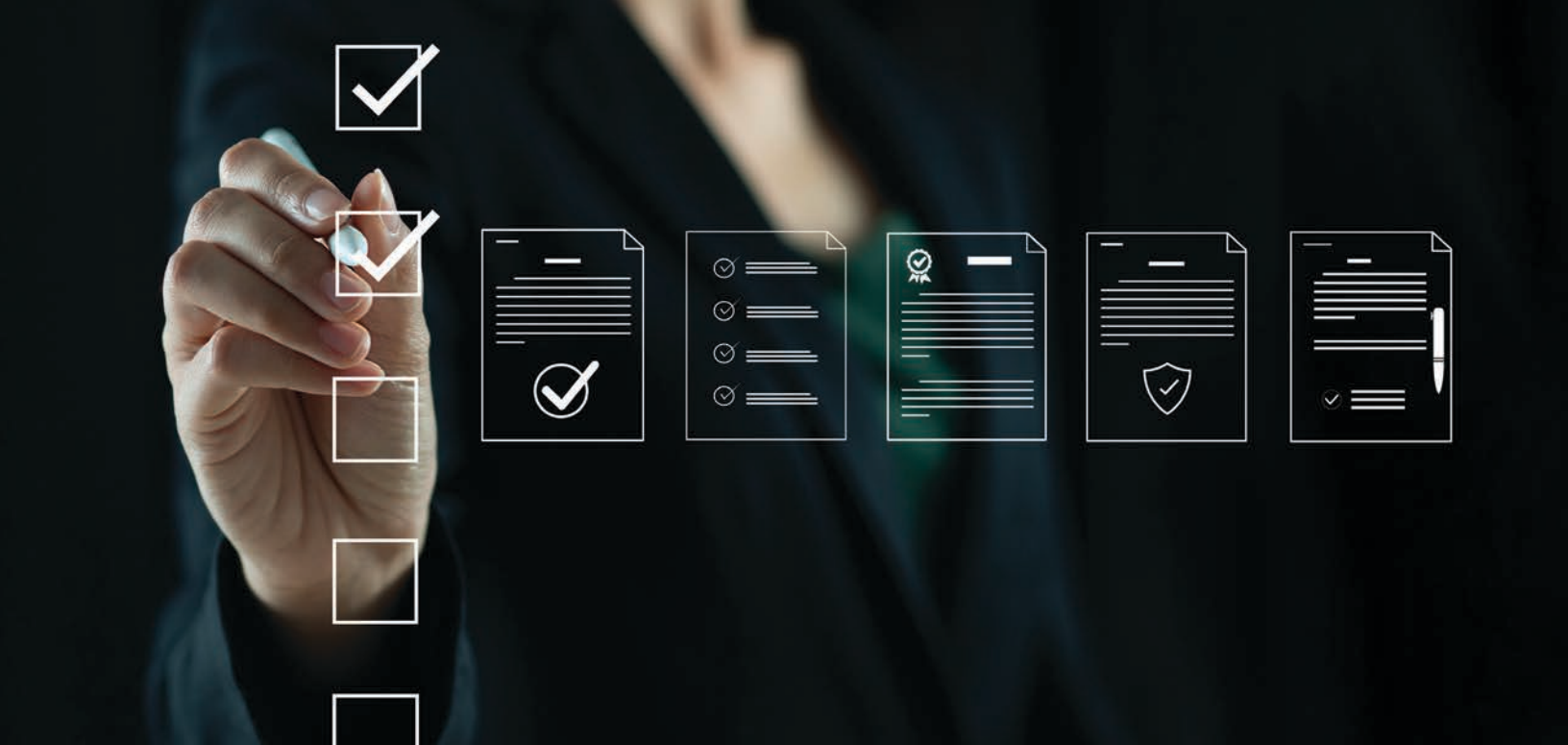
■ Why It Matters

Relying on vendors to submit notes generated offline and vendors' self-attestations may result in a risk that the outputs do not reflect the AI system's actual capabilities, limitations or performance in real-time. This provides limited assurance for Ontarians that the AI systems will work reliably in real-life clinical situations and generate consistently accurate medical notes.

Recommendation 9

For all future procurements related to AI systems, we recommend that Supply Ontario require mandatory live demonstrations, to assess the effectiveness of the AI systems.

For the auditee responses, see **Recommendations and Auditee Responses**.



4.4 The OPS's AI Strategy and Framework

4.4.1 The OPS's AI Strategy Lacks Many Key Components

The Ministry, on behalf of the OPS, developed an AI Strategy effective November 2024, ahead of similar initiatives in many other provinces and territories in Canada. As of September 2025, except for Quebec and the federal government, no other Canadian jurisdiction had a dedicated and publicly available AI strategy for its respective public sector.

The AI Strategy was developed ahead of similar initiatives in many other provinces and territories

We reviewed relevant white papers, academic research, best practices in other organizations, and Canadian and international public sector jurisdictional analyses, as well as selected private-sector AI strategy documents, based on publicly available information, to inform our understanding of the key elements that constitute a comprehensive and robust AI strategy.

When benchmarked against these best practices, we noted that the OPS AI strategy lacked several key components, such as:

- » specific actionable items;
- » identification of key sectors for AI use;
- » prohibition on use of AI in areas that pose unacceptable risks to the public;
- » an overarching investment commitment; and
- » environmental considerations.

Figure 8 outlines key components included in, and missing from, Ontario’s AI Strategy.

Figure 8: Inclusion of Key Components in OPS’s Public-Sector AI Strategy

Prepared by the Office of the Auditor General of Ontario

Components	Ontario
Specific actionable items	
» Objectives and overarching goals	✓
» Detailed initiatives to achieve the goals	✗
» Roles and responsibilities	✓
» Milestones and timelines for initiatives	✗
» Defined measurable outcomes for objectives, goals and initiatives	✗
» Metrics and targets for measurement of success of strategy	✗
Identification of key sectors for AI use	✗
Prohibition on use of AI in areas that pose unacceptable risks to the public	✗
Overarching investment commitment	✗
Environmental considerations	✗

Specific Actionable Items

The Ministry identified specific actionable items, such as objectives and overarching goals and defined roles and responsibilities. It did not include detailed initiatives, milestones and timelines, measurable outcomes, and performance metrics in its AI Strategy.

In comparison, the Government of Canada’s AI Strategy for the Federal Public Service 2025–2027 also outlines four priority areas. These priority areas have a set of accompanying actions that are “concrete, can be achieved or initiated within the AI Strategy’s two-year timeline, and have the greatest potential to advance responsible AI adoption within the public service,” as mentioned in the Government of Canada’s AI strategy published in March 2025.

OPS’s AI Strategy lacked many details for specific actionable items

The United Kingdom’s National AI Strategy, launched in 2021, is a 10-year vision with three-year rolling plans. These rolling plans have specific activities, milestones and timelines that support the achievement of the main priority areas of the AI strategy.

Identification of Key Sectors for AI Use

We noted that there was no measured and planned approach to identify and prioritize the use of AI within various sectors or program areas in the OPS. For example, we found that the Ministry had not identified key government sectors that could benefit from AI use by researching AI projects implemented in other jurisdictions.

The Ministry had not developed and maintained a structured pipeline of AI initiatives, categorized by short-, medium- and long-term implementation plans, nor had it assigned defined timelines to each project to support strategic planning and effective execution.

In comparison, we noted that the United States and Quebec identified potential AI use cases in the health, energy, education, transportation, defence and environment sectors as part of their AI strategies.

Prohibition on the Use of AI in Areas That Pose Unacceptable Risks to the Public

The Ministry had not explicitly identified any prohibited AI practices or areas where the use of AI should be banned in public service operations or to provide services to the public even though Ontario's *Enhancing Digital Security and Trust Act, 2024* allows for the drafting of regulations to prohibit AI use.

In comparison, the European Union's *Artificial Intelligence Act, 2024*, identifies the following unacceptable risks of AI use:

- » social scoring by governments, such as an AI-powered system that assigns a numerical score to a person's or group's behaviour or characteristics, where a good score can grant access to benefits like loans, while a bad score can result in restrictions on accessing goods and services;
- » AI that could manipulate human behaviour to cause harm;
- » real-time remote biometric identification in public spaces by law enforcement, with some narrow exceptions; and
- » AI systems that exploit vulnerabilities of children or persons with disabilities.

Overarching Investment Commitment

At the time of our audit, no standalone investment had been earmarked for the overall OPS AI Strategy.

An investment commitment toward the AI Strategy would demonstrate leadership engagement and support, as well as enable senior leadership to prioritize, set plans and manage approved budgets. It could encourage leaders, middle management and staff across the OPS to think about innovative solutions using AI for achieving intended outcomes.

In comparison, the Government of Canada, in its Budget 2024, announced dedicated investment in AI over five years for the public sector, starting in 2024/25. The United Kingdom earmarked investments from 2025 to 2030 to fund a national AI computer ecosystem as part of its National AI Strategy.

Environmental Considerations

We noted that the OPS's AI Strategy lacks considerations to assess the environmental impacts of using and procuring AI systems. We noted that there is no requirement in the AI Directive for ministries or the AI Executive Steering Committee to evaluate energy use, carbon emissions or green procurement criteria when adopting AI systems. This is a requirement in the strategies of all the other jurisdictions we reviewed.

Studies show that training one large AI model can emit over 280,000 kilograms of CO₂, comparable to the emissions of 60 gas-powered cars over a year.

Other jurisdictions are advancing efforts to assess the environmental impacts of AI. For example, the Government of Canada incorporated environmental and sustainability considerations as part of its AI initiatives. Departments deploying AI must assess environmental impacts as part of their strategic planning and risk management.

Similarly, the National Institute of Standards and Technology's NIST AI Risk Management Framework, a widely accepted industry standard, calls for evaluating and documenting the environmental and sustainability impacts of AI systems.

■ Why It Matters

Ontario can adopt best practices from other jurisdictions for its AI Strategy. In doing so it can be a leader in realizing the benefits of AI for the people of Ontario in a safe, efficient, effective and economic way. Without these best practices, there is a risk that effective AI adoption in the OPS, including operational efficiencies and service improvements, could be limited.

Recommendation 10

We recommend that the Ministry:

- research industry standards and AI strategic documents in other jurisdictions to review the key foundational elements supporting the achievement of AI strategic priorities;
- assess the applicability of these foundational elements with respect to an AI strategy for the OPS; and
- incorporate similar elements to enhance the robustness of its AI Strategy and support the timely achievement of its AI strategic priority areas.

For the auditee responses, see **Recommendations and Auditee Responses**.

Recommendations and Auditee Responses

For Recommendations 1 to 4, the Ministry of Public and Business Service Delivery and Procurement has commented on the progress it has made since we concluded our field work in November 2025. We have not audited or verified the information provided at this time and will follow up on it in our two-year follow-up report.

Recommendation 1

We recommend that the Cyber Security Division of the Ministry:

- review and block OPS staff's access to unsafe and unsecured AI websites; and
- implement measures to ensure that all staff take the training on AI-related risks with refresher training offered as the OPS advances its AI use.

Ministry Response

The Ministry agrees with the recommendation and the importance of ensuring the safety and security of data while using AI websites. The Ministry is working with key stakeholders to:

- assess the risk and business justification on the use of inappropriate/unsanctioned AI websites and implement required measures, including restricting access; and
- determine appropriate measures to ensure all staff take the training on AI-related risks, along with any required refresher training, and monitor compliance.

Since the audit concluded, the Ministry has blocked high-risk AI websites. The Ministry is operationalizing a process to perform regular risk assessments to identify additional candidates for blocking.

Recommendation 2

We recommend that the Ministry:

- establish KPI targets to measure and track Microsoft Copilot Chat's adoption;
- take actions to increase use of Microsoft Copilot Chat to the targeted rates and usage in the OPS; and
- report these KPIs to management on a monthly basis.

Ministry Response

The Ministry agrees with the recommendation to set clear targets, along with monitoring adoption, to support the effective use of Microsoft Copilot Chat across the Ontario Public Service.

The Ministry has created KPIs and is working with key stakeholders to:

- monitor and measure adoption on a routine basis as required by management;
- identify improvement opportunities; and
- develop action plans to further increase the use of Microsoft Copilot Chat.

Since the conclusion of the Auditor General’s fieldwork in November 2025, the Ministry has taken additional steps to strengthen adoption and awareness of Microsoft Copilot Chat through continued education sessions. This has helped increase usage to approximately 23,000 unique monthly users across the Ontario Public Service.

Recommendation 3

We recommend that the Cyber Security Division of the Ministry:

- block access to using Microsoft Copilot Chat on other browsers; and
- educate OPS staff through AI training about the dangers of using non-Microsoft browsers when accessing AI websites.

Ministry Response

The Ministry agrees with the recommendation and the importance of managing access.

The Ministry is evaluating options to mitigate this risk, including the restriction of access when warranted and within the technology’s capabilities.

The Ministry will continue to strengthen awareness and compliance by ensuring OPS staff receive ongoing AI-related training that emphasizes the security and privacy risks associated with using non-Microsoft browsers to access AI-enabled websites, and reinforces appropriate, approved and enterprise-sanctioned and supported AI-enabled websites.

Recommendation 4

We recommend that the Ministry validate test results provided by AI system vendors to ensure the testing was performed using a sufficient sample set representing Ontario’s demographics and that the testing meets all of the AI Directive requirements.

Ministry Response

The Ministry agrees with the recommendation of ensuring that testing of vendor-supported AI systems deployed across the OPS is robust, representative of and fully aligned with Ontario's Responsible Use of Artificial Intelligence Directive (AI Directive).

The Ministry is continuing to work with key partners to:

- further strengthen oversight of vendor-provided AI, including formalizing expectations for testing practices, validation approaches, transparency and documentation standards in alignment with the AI Directive;
- enhance assurance and evaluation practices to ensure testing and monitoring approaches reflect Ontario's demographic diversity where applicable, supports fairness and maintains clear accountability across the AI system life cycle; and
- advance procurement, governance and assurance mechanisms to further formalize review of vendor testing methodologies and confirm compliance with AI Directive requirements.

Since the conclusion of audit fieldwork, implementation of the Document Verification Service (DVS) included a phased-implementation plan that adopts diverse user-based testing and provided alternate channels such as in-person for users to access services.

Also, the Ministry has strengthened its approach to validating AI system behaviour and mitigating bias. To date, the Ministry has implemented a risk-based, layered approach that reduces reliance on vendor testing alone by grounding AI systems in curated, OPS-approved data sources and enforcing strict system behaviour controls.

Recommendation 5

For all future procurements related to AI systems, we recommend that Supply Ontario:

- increase the weighting assigned to criteria related to security and privacy controls, such as TRAs, PIAs and SOC reports, bias, and accuracy; and
- assign a minimum passing score or thresholds for these criteria.

Supply Ontario Response

Supply Ontario disagrees with the first action item of this recommendation. The combined weight between mandatory criteria and technical criteria is appropriate for security and privacy controls, bias and accuracy.

Supply Ontario agrees with the second action item of this recommendation and will include minimum thresholds in the technical evaluation stage for future AI-related systems to ensure security and privacy controls are emphasized.

Recommendation 6

We recommend that Supply Ontario:

- review industry standards and guidelines for other jurisdictions related to AI Scribe systems and implement best practices in Ontario; and
- require the AI Scribe system vendors to implement IT controls within their AI Scribe systems to enforce an attestation from users of the AI Scribe to confirm their review of notes produced.

Supply Ontario Response

Supply Ontario agrees with this recommendation. It will:

- review guidelines issued by the Information and Privacy Commissioner of Ontario and implement best practices for procurement for AI Scribe systems in Ontario; and
- work with vendors to determine the feasibility of including mandatory confirmation of notes review in future AI Scribe procurements.

Recommendation 7

We recommend that Supply Ontario, in collaboration with the Ministry:

- as part of its continuous vendor management process, obtain third-party reports, such as SOC 2 Type 2 reports and security assessment reports, from all AI Scribe system vendors annually;
- when procuring software of any type, including AI, always require a SOC 2 Type 2 report and/or security assessment to be included in the vendor response for assessment; and
- ensure the evaluator assesses whether the security assessment reports included the expected controls and review any exceptions in the reports for potential risks.

Supply Ontario Response

Supply Ontario agrees with this recommendation and has included the annual provision of SOC 2 Type 2 reports in all current AI Scribe contracts. Supply Ontario will assess mandatory provision of SOC 2 Type 2 reports and/or security assessments as part of vendor responses where feasible and will ensure for evaluations fully assess the reports provided for future software procurements.

Recommendation 8

For all future procurements related to AI systems, we recommend that Supply Ontario follow the principles of the AI Directive and evaluate bias risks related to AI systems by requiring evidence of bias testing results or doing their own bias testing prior to selecting a system.

Supply Ontario Response

Supply Ontario agrees with this recommendation and will require bias testing results to be provided by vendors for evaluation for future procurements related to AI systems.

Recommendation 9

For all future procurements related to AI systems, we recommend that Supply Ontario require mandatory live demonstrations, to assess the effectiveness of the AI systems.

Supply Ontario Response

Supply Ontario agrees with this recommendation and will consider the use of live demonstrations as part of the evaluation or award process for future AI system procurements.

Recommendation 10

We recommend that the Ministry:

- research industry standards and AI strategic documents in other jurisdictions to review the key foundational elements supporting the achievement of AI strategic priorities;
- assess the applicability of these foundational elements with respect to an AI strategy for the OPS; and
- incorporate similar elements to enhance the robustness of its AI Strategy and support the timely achievement of its AI strategic priority areas.

Ministry Response

The Ministry agrees with the recommendation to further strengthen the OPS AI strategy. Setting strategic priorities for adopting a fast-moving technology requires regular, iterative reviews.

The early OPS AI strategy reviewed by the Auditor General focused efforts on the technology supports necessary to responsibly operationalize AI. Future versions of the strategy will focus on ensuring that business adoption is advancing within the appropriate ethical, security and privacy controls.

To address this recommendation, the Ministry will:

- ensure the work is informed by industry and cross-jurisdictional insights;
- include relevant standards, governance and accountability structures, with clear priorities and action items that support responsible AI adoption and use within the OPS; and
- set a clear timeline for regular reviews.

The Ministry will continue to work collaboratively across the enterprise to ensure the OPS AI Strategy is coherent, informed by evidence and appropriate to the needs of Ontario.

Audit Criteria

In planning our work, we identified the audit criteria we would use to address our audit objective (outlined in **Section 3**). These criteria were established based on a review of applicable legislation, policies and procedures, internal and external studies, and best practices. Senior management at the Ministry reviewed and acknowledged the suitability of our objective and associated criteria:

1. A best-practice AI strategy is developed and communicated internally as well as publicly.
2. A comprehensive framework is developed to execute the AI strategy across the OPS.
3. A clearly defined governance structure is established to monitor compliance with the AI Framework and AI Directive for use cases, and corrective actions are taken for non-compliance.
4. Best-practice key performance indicators, including targets, are established to measure the success of AI adoption. Deviations from established targets are addressed and reported to the Ministry and/or Treasury Board Secretariat.
5. The Ministry considers the following, at a minimum, while identifying and selecting the most appropriate AI tools and technologies:
 - a. basic principles of cybersecurity such as data residency, tenant and data restriction, inappropriate access, data security, and privacy; and
 - b. having due regard for economy.
6. Mitigating controls, in accordance with industry standards, are established to address the risks associated with implementing/using vulnerable AI tools and technologies.
7. Mitigating controls, in accordance with industry standards, are established to address the risks relating to cybersecurity, data residency, data governance, privacy and third-party vendors.
8. The Ministry has conducted an assessment to identify its talent gaps and resource gaps with respect to AI technologies.
9. Relevant training is provided to OPS staff to educate them on safe use of AI tools and to make them aware of risks and threats of AI tools.

Audit Approach

We conducted our audit between January 2025 and November 2025. We obtained written representation from the Ministry's management that, effective April 28, 2026, they had provided us with all the information they were aware of that could significantly affect the findings or the conclusion of this report.

As part of our audit work, we:

- » interviewed relevant staff from the Ministry, the Office of the Corporate Chief Information Officer (including the Government Services Integrated Cluster, the Cyber Security Division, the Enterprise Digital & Technology Strategy Division and the Supply Chain Division), the Treasury Board Secretariat and Supply Ontario;
- » interviewed members of the AI Executive Steering Committee to understand the AI project review and endorsement process;
- » reviewed key strategic documents, including the AI Strategy, AI Framework, AI Directive, AI Playbook and AI Workbook, for consistency and to ensure they included key foundational elements consistent with best practices found in other jurisdictions;
- » reviewed and performed detailed testing on a selection of AI projects to test their compliance with the requirements outlined in the AI Directive, for AI projects that are fully implemented, initiated or operating in pilot or "proof of concept" phase, for the last five years (2020/21 to 2024/25);
- » reviewed cybersecurity assessments conducted for AI projects that are fully implemented;
- » reviewed relevant procurement documents, including the request for bids, vendors' bids, evidentiary documents, vendors' assessments and guidelines, to evaluate whether key risks and ethical principles were considered while procuring AI systems from third-party vendors; and
- » reviewed performance metrics, including usage metrics and statistics, as well as targets established to monitor the success of the OPS AI Strategy and the adoption of AI systems that are fully implemented across the OPS.

Audit Opinion

To the Honourable Speaker of the Legislative Assembly:

We conducted our work for this audit and reported on the results of our examination in accordance with Canadian Standard on Assurance Engagements 3001—*Direct Engagements* issued by the Auditing and Assurance Standards Board of the Chartered Professional Accountants of Canada. This included obtaining a reasonable level of assurance.

The Office of the Auditor General of Ontario applies Canadian Standards on Quality Management and, as a result, maintains a comprehensive system of quality management that includes documented policies and procedures with respect to compliance with rules of professional conduct, professional standards and applicable legal and regulatory requirements.

We have complied with the independence and other ethical requirements of the Code of Professional Conduct of the Chartered Professional Accountants of Ontario, which are founded on fundamental principles of integrity, objectivity, professional competence and due care, confidentiality and professional behaviour.

We believe the audit evidence we have obtained is sufficient and appropriate to provide a basis for our conclusions.

May 12, 2026



Shelley Spence, FCPA, FCA, LPA

Auditor General
Toronto, Ontario

Acronyms

Acronym	Definition
AI	artificial intelligence
DVS	Document Verification Service
GenAI	generative artificial intelligence
ISO 27001	International Organization for Standardization 27001
IT	information technology
KPI	key performance indicator
LLM	large language model
OPS	Ontario Public Service
PIA	privacy impact assessment
RFB	request for bids
SOC	System and Organization Controls
TRA	threat risk assessment
VOR	vendor of record

Glossary

Term	Definition
AI Directive	A policy that sets guidelines for responsible and ethical deployment of AI. It helps to ensure that AI is used in a way that is transparent, responsible and accountable.
AI Executive Steering Committee	A strategic group that guides the direction, governance and implementation of AI within the Ministry.
AI model	A program that has been trained on a set of data to recognize certain patterns or make certain decisions without further human intervention.
AI Playbook	A strategic guide that provides organizations with a framework, best practices and step-by-step processes for implementing AI initiatives successfully.
AI system	A software system that leverages AI to solve a specific business or operational problem.
AI Workbook	A guidance tool used to support the implementation of the AI Directive. It focuses on planning, evaluating and explaining AI systems in alignment with ethical and legal standards.
Artificial intelligence (AI)	The ability of IT systems to perform tasks that typically require human intelligence, such as automating tasks, analyzing large complex data for decision-making and problem-solving based on the inputs it is provided with.
Bias	An action or attitude of supporting or opposing a particular person or thing in an unfair way.
Biometric identification	The process of verifying a person's identity by using unique or physical characteristics like fingerprints, facial features or voice patterns.
Data residency	The physical or geographic location where data is stored and maintained.
GenAI	A subset of AI that emulates various aspects of human intelligence, generating texts, images, videos, music and software code from prompts.
Hallucinations	AI hallucinations occur when AI systems generate outputs that are made up, fabricated or not based on the actual data provided to them.
Large language model (LLM)	A type of AI program that uses deep learning to analyze and generate human-like text by learning patterns and relationships from massive datasets of text and code.
Microsoft Copilot Chat	A conversational, AI-powered website that boosts productivity by providing contextual help, automating routine tasks, and analyzing data to support decision-making.
Privacy risk assessment (PIA)	Identifies and evaluates potential privacy risks associated with an organization's collection, use and disclosure of personal information, assessing their likelihood and impact to help inform decision-making for managing risks and implementing controls to prevent or at least reduce privacy breaches.

Term	Definition
System and Organization Controls (SOC) report	A third-party audit report that explains how well a vendor's processes are designed and whether they work as intended to ensure confidentiality, integrity, security, and availability of data and operations.
Threat risk assessment (TRA)	Identifies and evaluates potential threats to an organization's information systems, assessing their likelihood and impact to inform decision-making for managing risks and implementing controls to help prevent or at least reduce cyberattacks.
Transcribe	To convert audio text into handwritten or typed text, or handwritten text into typed text.
Use case	An AI system developed to address a particular problem or address a specific need.

Appendix 1: Key AI Systems in the OPS, September 2025

Source of data: Ministry of Public and Business Service Delivery and Procurement

AI System	Ministry	Description	Vendor	Go Live
Copilot Chat	OPS-wide	Copilot Chat was made available to all OPS staff to research, summarize information, generate content and write software code.	Microsoft	May 2024
Copilot M365	Limited deployment (500)	Copilot for Microsoft 365 is a GenAI-powered assistant integrated within Microsoft Office applications such as Word, Excel, Outlook and Teams. It uses LLMs to help OPS staff work more efficiently by generating content, summarizing information and providing insights.	Microsoft	Pilot – Oct 2024; wide release – Dec 2025
Cyber Control Tower	Service Ontario – MPBSDP	An advanced AI chatbot system that will enable staff to get quick insights into driver and vehicle transaction data based on questions asked by the staff.	Amazon, Microsoft and Google	Dec 2025
Document Verification Service (DVS)	MPBSDP	The DVS system will allow Ontarians to verify their identities online to access various government services.	Interac	Not announced
Position Information Questionnaire	Treasury Board Secretariat	This project will leverage AI to summarize survey responses from a large number of OPS staff, organizing the insights by job type.	Microsoft	Oct 2026
Red Tape Reduction Regulatory Modernization (REGi)	Ministry of Red Tape Reduction	A system to help OPS staff from various ministries quickly go through large volumes of laws, rules and regulations, and identify requirements laid out in these sources.	Google Gemini	Jun 2025

Appendix 2: Criteria Used to Evaluate Bidders in the RFB Process for AI Scribe Systems

Source of data: Supply Ontario

Criteria	Description	Max Score	# of Requirements
Domestic presence in Ontario	Evaluation of the vendor's presence in the province based on questions related to local partnerships, provincial experience, ownership and corporate presence in Ontario and the location of their headquarters.	159	7
Clinical format, including transcription of discussions between multiple speakers	Evaluation of whether the vendor's AI Scribe system offers a variety of templates for medical note generation, whether it can be customized and modified in-session, and whether it has the capability to support multiple conversations and additional languages. The requirements also assess whether the vendor continuously monitors and enhances the accuracy and reliability of notes generated by their AI Scribe systems, as technology evolves.	47	11
Accuracy of medical notes generated	This section validates the system's functionality and performance, by requiring vendors to provide an unedited system-generated medical note based on simulated recordings of conversations between health-care professionals and patients.	20	1
Business, including system interface, vendor onboarding, training and pricing	Evaluation of whether the system is designed with a simple intuitive interface, and offers compatibility across multiple platforms and devices, as well as connectivity. The evaluation also covered the vendor's onboarding and training program, and pricing options.	28	6
Data privacy/legal controls	Evaluation of the vendor organization's adherence to storing and processing personal information exclusively in Canada. Additionally, this section also assesses the system's risk management and incident management program for privacy breaches and security incidents, its enterprise business continuity framework, the adequacy of its security and privacy training, and its compliance with the principles outlined in the <i>Personal Information Protection and Electronic Documents Act, 2000</i> and the <i>Personal Health Information Protection Act, 2004</i> .	130	10

Criteria	Description	Max Score	# of Requirements
System security controls	Evaluation of the vendor organization's security controls, including access management, automated backups, disaster recovery, encryption and audit logging mechanisms. It also reviews the organization's risk assessment processes (e.g., TRAs and PIAs); external auditor reports, such as SOC 2 Type 2 reports; and security certifications, including ISO 27001 and HITRUST r2.	93	15
Contract negotiation, including vendor-requested amendments	This section evaluates whether the vendor is willing to accept Supply Ontario's approved vendor contract without modifications. Any revisions to the standard contract that are requested by the vendor results in reducing the total number of points assigned to this criterion.	53	1
Total		530	51



© 2026, King's Printer for Ontario
ISBN 978-1-4868-9746-9 (PDF)

An electronic version of this report is available at www.auditor.on.ca
Ce document est également disponible en français.

Cover photograph credit: ©iStockphoto.com