TREASURY INSPECTOR GENERAL FOR TAX ADMINISTRATION

Assessment of Artificial Intelligence Risk Management Practices

September 24, 2026

Report Number 2026-208-059

TIGTACommunications@tigta.treas.gov | www.tigta.gov

HIGHLIGHTS: Assessment of Artificial Intelligence Risk Management Practices

Final Audit Report issued on September 24, 2026 **Report Number 2026-208-059**

Why TIGTA Did This Audit

Office of Management and Budget (OMB) Memorandum M-25-21 established new Artificial Intelligence (AI) governance requirements with phased implementation deadlines. Memorandum M-25-21 required federal agencies, (*i.e.*, the Department of the Treasury) to implement minimum risk management practices to manage the risks from high-impact AI use cases by April 2026.

The National Institute of Standards and Technology outlines specific challenges managing AI risks and its dependence on large scale datasets, which can impact data quality and raise validity concerns. Furthermore, policies and resources should be prioritized based on the assessed risk level and potential impact of an AI system.

We evaluated the IRS's process to document its data reliability assessments for its AI use cases.

Impact on Tax Administration

High-impact AI is any AI system whose output serves as a principal basis for decisions or actions. These decisions or actions have a legal, material, binding, or significant effect on civil rights, privacy, health, safety, or access to critical government resources and services.

AI risk management for high-impact use cases is critical because these use cases may have a material effect on taxpayer outcomes. The responsible use of AI depends on data that are appropriate for the AI modeling approach and use.

What TIGTA Found

The IRS has been using AI for many years. For example, we previously reported that in 2017, the IRS undertook initiatives to improve tax return classification and issue selection using AI. As of December 2025, the IRS had an inventory of 225 AI use cases.

We reviewed five AI use cases that the IRS presumed to be high-impact and found that risk management practices need improvement. Specifically, the IRS did not fully document:

- Its evaluation of risk for its AI use cases. Our review found that 40 percent (2 of 5) of AI use cases did not have an impact assessment documented. The remaining 60 percent (3 of 5) of AI use cases did have a documented impact assessment. As of July 2026, the IRS has not received any guidance or instructions from the Department of the Treasury to implement minimum risk management practices. The IRS developed and began implementing its own AI governance policy, risk management processes, and supporting documentation to align with OMB Memorandum M-25-21.
- How it evaluated the quality and appropriateness of the data used in its AI use cases. According to the program managers and data scientists we interviewed, the data quality checks were completed; however, the documentation for each use case varied. For example, 20 percent (1 of 5) of the cases had documents outlining numerous processes and procedures used in testing. The remaining 80 percent (4 of 5) did not have any testing documentation. Research, Applied Analytics and Statistics management agreed that documentation enables repeatability and consistency and can enhance risk management decisions. However, they did not prioritize documenting or standardizing baseline procedures or guidelines for evaluating data quality checks before their use in high-impact AI models.

What TIGTA Recommended

We made two recommendations to the IRS's Chief Data and Analytics Officer to ensure that: AI impact assessments for high-impact use cases are completed in accordance with OMB Memorandum M-25-21; and to develop and implement standard processes for evaluating data quality for use in AI use cases.

The IRS agreed with both recommendations and stated that it either has or plans to implement corrective actions.

TREASURY INSPECTOR GENERAL
FOR TAX ADMINISTRATION

U.S. DEPARTMENT OF THE TREASURY

WASHINGTON, D.C. 20024

September 24, 2026

MEMORANDUM FOR: COMMISSIONER OF INTERNAL REVENUE

Diana M. Tengesdal

FROM: Diana M. Tengesdal
Deputy Inspector General for Audit

SUBJECT: Final Audit Report – Assessment of Artificial Intelligence Risk Management Practices (Audit No.: 2026208011)

This report presents the results of our review to determine whether the data within the Internal Revenue Service's (IRS) Compliance Data Warehouse is sufficiently reliable to support the use of Artificial Intelligence (AI). While we did not evaluate the accuracy of the data itself, our review focused on the IRS's process to document its data reliability assessments for its AI use cases. This review is part of our Fiscal Year 2026 Annual Program Plan and addresses the major management and performance challenge of *Modernizing Information Technology*.

Management's complete response to the draft report is included in Appendix II. If you have any questions, please contact me or Linna K. Hung, Assistant Inspector General for Audit (Security and Information Technology Services).

Table of Contents

Background

The federal government recognizes the potential of Artificial Intelligence (AI) to increase efficiency and improve government services. Key AI terminologies are defined as follows:

- **AI** – A machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments.
- **AI model** – A component of an information system that implements AI technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs.
- **AI use case** – The specific scenario that AI is designed, developed, procured, or used to advance the execution of agencies' missions and their delivery of programs and services, enhance decision-making, or provide the public with a particular benefit.

Executive Order 13960 encouraged federal agencies to continue to use AI, when appropriate, to benefit the American people.[1] Agencies must therefore design, develop, acquire, and use AI in a manner that fosters public trust and confidence. Additionally, agencies shall be transparent in disclosing relevant information regarding their use of AI to appropriate stakeholders, including Congress and the public, to the extent practicable and in accordance with applicable laws and policies. Other executive orders addressing AI have been issued, including an executive order issued in January 2025, which addresses the topic of removing barriers to American leadership in the development and use of AI.[2]

The Internal Revenue Service (IRS) has been using AI for many years. For example, we previously reported that in 2017, the IRS undertook initiatives to improve tax return classification and issue selection using AI.[3] As of December 2025, the IRS had an inventory of 225 AI use cases.

The IRS uses different data platforms to train or use AI use cases, such as the Selection and Analytic Platform, the Enterprise Data Platform, and the Compliance Data Warehouse (hereafter referred to as data warehouse). The Research, Applied Analytics, and Statistics (RAAS) organization under the direction of IRS's Chief Data and Analytics Officer serves as the system owner of the data warehouse and manages the IRS-wide AI use case inventory.

The Office of Management and Budget (OMB) issued a memorandum in April 2025, which established new AI governance requirements as well as phased implementation deadlines

[1] Exec. Order No. 13960, *Promoting the Use of Trustworthy Artificial Intelligence in the Federal* Government, 85 Fed. Reg. 78939 (2020). This order expanded upon Exec. Order No. 13859, *Maintaining American Leadership in Artificial Intelligence*, 84 Fed. Reg. 3967 (2019).

[2] Exec. Order No. 14179, *Removing Barriers to American Leadership in Artificial Intelligence*, 90 Fed. Reg. 8741 (2025).

[3] Treasury Inspector General for Tax Administration, Report No. 2025-308-022, *The IRS Could Leverage Examination Results in Artificial Intelligence Examination Case Selection Models and Improve Processes to Evaluate Performance*, May 2025.

through April 2026.[4] This memorandum also revoked previous guidance outlined in OMB Memorandum M-24-10.[5]

Results of Review

Artificial Intelligence Risk Management Practices Need Improvement

We reviewed five of the AI use cases that the IRS presumed to meet the definition of high-impact.[6] High-impact AI is any AI system whose output serves as a principal basis for decisions or actions.[7] These decisions or actions have a legal, material, binding, or significant effect on civil rights, privacy, health, safety, or access to critical government resources and services. The five AI use cases we reviewed used data from the IRS's data warehouse. Our review found that the IRS did not fully document its evaluation of risk and how it validated data for its AI use cases.

Evaluation of risk for AI use cases was not fully documented

The IRS did not fully document its evaluation of risk for AI use cases. OMB Memorandum M-25-21 requires that agencies implement minimum risk management practices for high-impact AI by April 2026. These risk management practices require a completed and documented impact assessment to include an assessment of the:

- Intended purpose for the AI and its expected benefit, supported by specific metrics or qualitative analysis.
- The quality and appropriateness of the relevant data and model capability, supported by a summary of the data used in AI's design, development, training, testing, and operation and its fitness for the AI's intended purpose.
- Potential impacts of AI use.

Furthermore, OMB Memorandum M-25-21 states that if a high-impact use case is not compliant with the minimum risk management practices, then the agency must safely discontinue use of the AI functionality.

Our review of 5 deployed high-impact AI use cases found that 40 percent (2 of 5) AI use cases did not have an impact assessment documented. The remaining 60 percent (3 of 5) of the AI use cases did have a documented impact assessment.[8]

---

[4] Office of Management and Budget, Memorandum M-25-21, *Accelerating Federal Use of AI through Innovation, Governance, and Public Trust* (April 2025).

[5] Office of Management and Budget, Memorandum M-24-10, *Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence* (March 2024).

[6] Certain uses of AI are presumed to be high-impact. AI may be integrated in decision or activity pipelines in high-impact categories without meeting the definition of high-impact when the AI's output does not actually "serve as a principal basis for" the relevant type of agency action or decision.

[7] See Appendix III for a glossary of terms.

[8] AI impact assessments were conducted prior to the effective date of OMB Memorandum M-25-21 and based on prior guidance.

RAAS management stated that they prioritized compliance with federal requirements for governance of AI use cases. However, OMB Memorandum M-25-21 required federal agencies (*i.e.*, the Department of the Treasury) to implement minimum risk management practices to manage the risks from high-impact AI use cases by April 2026. In February 2026, the IRS developed and began implementing its own AI governance policy, risk management processes, and supporting documentation to align with OMB Memorandum M-25-21. RAAS management stated that as of July 2026, they have not received any guidance or instructions from the Department of the Treasury to implement minimum risk management practices.

The responsible use of AI depends on data that are appropriate for the AI modeling approach and use. Therefore, AI risk management for high-impact use cases is critical because these use cases may have a material effect on taxpayer outcomes. For example, according to the IRS's data warehouse internal web page, the data from the warehouse should not be used as a substitute for source systems for operational decision-making purposes. The IRS's impact assessments should include documented analysis of the quality and appropriateness of the data from the data warehouse, its potential impact on taxpayers, and planned mitigation measures. The absence of consistent and formal documentation of the risks and impacts of developed systems can lead to oversights in the application of risk mitigation strategies.

Fitness testing and documentation was inconsistent

The IRS did not fully document how it evaluated the quality and appropriateness of the data used in its AI use cases in the AI impact assessments. In addition, the documentation available was different for each AI use case we reviewed. OMB Circular A-123 states that proper documentation is a necessary part of an effective internal control system.[9] Documentation preserves evidence to support decision-making.

We interviewed IRS program managers and data scientists to discuss steps taken to determine whether the data were "fit" for use (*i.e.*, the quality and appropriateness of the data). According to them, data quality checks were completed, and the data were fit for use. However, the type of data quality checks and documentation completed for each AI use case varied. For example, 20 percent (1 of 5) of the use cases we reviewed had documents outlining numerous processes and procedures used in testing. The remaining 80 percent (4 of 5) did not have any testing documentation. In addition, the documentation outlined detailed information related to change management (*i.e.*, any modifications to the model, such as updates to scoring, methodologies, or execution processes that are formally reviewed and approved) oversight, and validation. Similar documents do not exist for other use cases that we evaluated.

RAAS management agreed that standard documentation procedures would be valuable. Comprehensive documentation enables repeatability and consistency and can enhance risk management decisions. In addition, documentation demonstrates consistent compliance with federal requirements for AI risk management. However, they did not prioritize documenting or standardizing baseline procedures or guidelines for evaluating data quality checks before its use in high-impact AI models.

The National Institute of Standards and Technology outlines specific challenges managing AI risks and its dependence on large scale datasets, which can impact data quality and validity

[9] Office of Management and Budget, Circular A-123, *Management's Responsibility for Internal Control* (March 2026).

concerns.[10] Furthermore, policies and resources should be prioritized based on the assessed risk level and potential impact of an AI system. Higher priority should be given:

- When the AI system is trained on large datasets comprised of sensitive or protected data such as Personally Identifiable Information (*e.g.*, Social Security Numbers).
- Where the outputs of the AI systems have an impact on humans.

A lack of formal written procedures or guidelines for assessing the quality and appropriateness of the data used in high-impact AI use cases may result in inaccurate, biased, or unreliable AI outputs that can cause significant financial losses and reputational damage. In addition, poor data quality may cause AI projects to underperform.

The Chief Data and Analytics Officer, should:

Recommendation 1: Ensure that AI impact assessments for high-impact use cases are completed in accordance with OMB Memorandum M-25-21 requirements.

> **Management's Response:** IRS management agreed with this recommendation and stated that they will complete AI impact assessments for currently deployed high-impact AI use cases as of November 2026.

Recommendation 2: Develop and implement standard processes for evaluating data quality and documenting the fitness of data for use in AI use cases.

> **Management's Response:** IRS management agreed with this recommendation and stated that they have implemented standard processes for evaluating data quality and documenting the fitness of data for use in AI use cases.

[10] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework 1.0* (January 2023).

Appendix I

Detailed Objective, Scope, and Methodology

The overall objective of this review was to determine whether the data within the IRS's Compliance Data Warehouse is sufficiently reliable to support the use of AI. While we did not evaluate the accuracy of the data itself; our review focused on the IRS's process to document its data reliability assessments for its AI use cases. To accomplish our objective, we:

- Assessed whether the IRS implemented the minimum risk management practices in a consistent, documented, and operational manner by reviewing federal guidance on high-impact use case requirements, conducted interviews with the assurance team. We identified seven deployed high-impact use cases using data from the IRS's data warehouse. We reviewed the five AI use cases used for workload selection.
- Determined whether the IRS identified and documented its AI use cases by interviewing the AI assurance team, reviewing the inventory, assessing federal guidance, validated reporting requirements, and compared submitted inventory data against parent agency records, assessed non-reportable cases for accuracy and confirmed that each use case contained all required reporting elements.

Performance of This Review

This review was performed with information obtained from the Research, Applied Analytics and Statistics organization located in Washington, D.C., during the period October 2025 through April 2026. We conducted this performance audit in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objective.

Internal Controls Methodology

Internal controls relate to management's plans, methods, and procedures used to meet their mission, goals, and objectives. Internal controls include the processes and procedures for planning, organizing, directing, and controlling program operations. They include the systems for measuring, reporting, and monitoring program performance. We determined that the following internal controls were relevant to our audit objective: Federal guidance from the OMB and the National Institute of Standards and Technology, as well as the Internal Revenue Manual. We evaluated these controls by interviewing personnel from the RAAS organization and reviewing AI impact assessment documentation.

Appendix II

Management's Response to the Draft Report

RESEARCH, APPLIED ANALYTICS AND STATISTICS

DEPARTMENT OF THE TREASURY
INTERNAL REVENUE SERVICE
WASHINGTON, DC 20224

September 21, 2026

MEMORANDUM FOR DIANA M. TENGESDAL
DEPUTY INSPECTOR GENERAL FOR AUDIT

FROM: Lucia Lykke Lucia C. Lykke Digitally signed by Lucia C. Lykke Date: 2026.09.21 14:18:17 -04'00'
Acting Chief Data and Analytics Officer

SUBJECT: Draft Audit Report – Assessment of Artificial Intelligence Risk Management Practices (Audit #2026208011)

Thank you for the opportunity to review your draft report entitled Assessment of Artificial Intelligence Risk Management Practices.

Your report evaluated the Internal Revenue Service's (IRS) procedures for documenting the risk management practices of high-impact artificial intelligence (AI) use cases and recommended that the IRS ensure compliance with Office of Management and Budget (OMB) M-25-21 requirements.

We remain committed to continuous improvement and transparency in our AI risk management practices at the IRS.

Our response to the Treasury Inspector General for Tax Administration's recommendations are attached. If you have any questions, please contact me at (901) 512-8724 or Lucia.C.Lykke@irs.gov.

Attachment

Assessment of Artificial Intelligence Risk Management Practices
(Audit No.: 2026208011)

Recommendations

RECOMMENDATION 1:
The Chief Data and Analytics Officer should ensure that Artificial Intelligence (AI) impact assessments for high-impact use cases are completed in accordance with OMB Memorandum M-25-21 requirements.

CORRECTIVE ACTION:
The Chief Data and Analytics Officer (CDAO) agrees to complete AI impact assessments for currently deployed high impact AI use cases as of November 2026.

IMPLEMENTATION DATE:
November 2026

RESPONSIBLE OFFICIAL:
Chief Data and Analytics Officer

RECOMMENDATION 2:
The Chief Data and Analytics Officer should develop and implement standard processes for evaluating data quality and documenting the fitness of data for use in AI use cases.

CORRECTIVE ACTION:
The CDAO implemented standard processes for evaluating data quality and documenting the fitness of data for use in AI use cases.

IMPLEMENTATION DATE:
July 13, 2026

RESPONSIBLE OFFICIAL:
Chief Data and Analytics Officer

Appendix III

Glossary of Terms

Term	Definition
Artificial Intelligence System	Any data system, software, application, tool, or utility that operates in whole or in part using dynamic or static machine-learning algorithms or other forms of AI, whether it is intended for research, development, implementation, or integration into agency systems or operational workflows.
Data Scientist	Data scientists are experts at extracting specific insights and answers from data. They have a deep understanding of the industry or business discipline that they work in, and they rely on widely used programming languages to conduct exploratory data analysis and statistical regression.
Data Warehouse	A repository of compliance, filing, and related databases that are used to support projects, analyses, and studies related to tax administration.

Appendix IV

Abbreviations

AI	Artificial Intelligence
IRS	Internal Revenue Service
OMB	Office of Management and Budget
RAAS	Research, Applied Analytics and Statistics
TIGTA	Treasury Inspector General for Tax Administration

To report fraud, waste, or abuse,
contact our hotline on the web at
https://www.tigta.gov/reportcrime-misconduct.

To make suggestions to improve IRS policies, processes, or systems
affecting taxpayers, contact us at
TIGTACommunications@tigta.treas.gov.

Information you provide is confidential, and you may remain anonymous.