GAO

ARTIFICIAL INTELLIGENCE

Fully Implementing Key Practices Could Help DHS Ensure Responsible Use for Cybersecurity

Accessible Version

Report to Congressional Addressees

GAO Highlights

February 2024

Artificial Intelligence

Fully Implementing Key Practices Could Help DHS Ensure Responsible Use for Cybersecurity

Why GAO Did This Study

Executive Order No. 14110, issued in October 2023, notes that while responsible AI use has the potential to help solve urgent challenges and make the world more secure, irresponsible use could exacerbate societal harms and pose risks to national security. Consistent with requirements of Executive Order No. 13960, issued in 2020, DHS has maintained an inventory of its AI use cases since 2022.

This report examines the extent to which DHS (1) verified the accuracy of its inventory of AI systems for cybersecurity and (2) incorporated selected practices from GAO's AI Accountability Framework to manage and oversee its use of AI for cybersecurity.

GAO reviewed relevant laws, OMB guidance, and agency documents, and interviewed DHS officials. GAO applied 11 key practices from the Framework to DHS's AI cybersecurity use case—Automated PII Detection. DHS uses this tool to prevent unnecessary sharing of PII. GAO selected the 11 key practices to reflect all four Framework principles, align with early stages of AI adoption, and be highly relevant to the specific use case.

What GAO Recommends

GAO is making eight recommendations to DHS, including that it (1) expand its review process to include steps to verify the accuracy of its AI inventory submissions, and (2) fully implement key AI Framework practices such as documenting sources and ensuring the reliability of the data used. DHS concurred with the eight recommendations.

What GAO Found

To promote transparency and inform the public about how artificial intelligence (AI) is being used, federal agencies are required by Executive Order No. 13960 to maintain an inventory of AI use cases. The Department of Homeland Security (DHS) has established such an inventory, which is posted on the Department's website.

However, DHS's inventory of AI systems for cybersecurity is not accurate. Specifically, the inventory identified two AI cybersecurity use cases, but officials told us one of these two was incorrectly characterized as AI. Although DHS has a process to review use cases before they are added to the AI inventory, the agency acknowledges that it does not confirm whether uses are correctly characterized as AI. Until it expands its process to include such determinations, DHS will be unable to ensure accurate use case reporting.

DHS has implemented some but not all of the key practices from GAO's AI Accountability Framework for managing and overseeing its use of AI for cybersecurity. GAO assessed the one remaining cybersecurity use case known as Automated Personally Identifiable Information (PII) Detection—against 11 AI practices selected from the Framework (see figure).

Status of the Department of Homeland Security's Implementation of Selected Key Practices to Manage and Oversee Artificial Intelligence for Cybersecurity
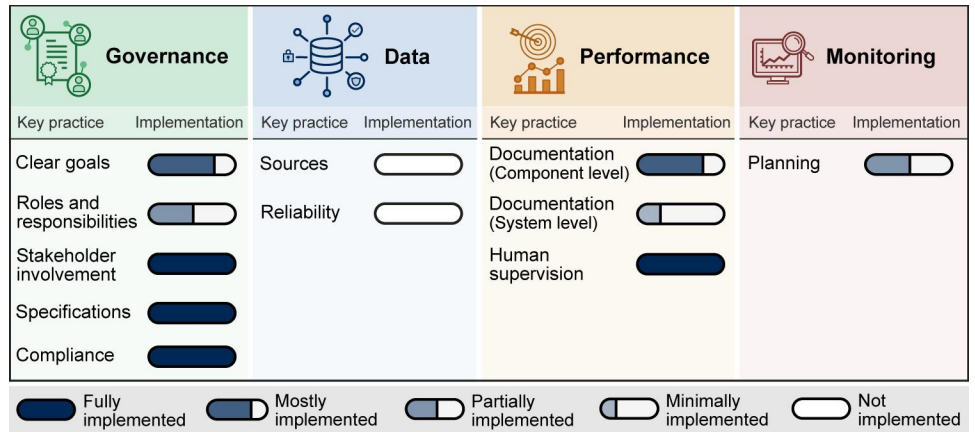
Source: GAO analysis of agency documents and interviews with Department of Homeland Security officials; GAO (icons). | GAO-24-106246

Accessible text for Status of the Department of Homeland Security's Implementation of Selected Key Practices to Manage and Oversee Artificial Intelligence for Cybersecurity

- Governance
    - Clear goals – Mostly implemented
    - Roles and responsibilities – Partially implemented
    - Stakeholder involvement – Fully implemented
    - Specifications – Fully implemented
    - Compliance – Fully implemented

GAO-24-106246 Artificial Intelligence

- Data
  - Sources – Not implemented
  - Reliability – Not implemented
- Performance
  - Documentation (component) – Mostly implemented
  - Documentation (system) – Partially implemented
  - Human supervision – Fully implemented
- Monitoring
  - Planning – Partially implemented

Source: GAO analysis of agency documents and interviews with DHS officials. | GAO-24-106246

GAO found that DHS fully implemented four of the 11 key practices and implemented five others to varying degrees in the areas of governance, performance, and monitoring. It did not implement two practices: documenting the sources and origins of data used to develop the PII detection capabilities, and assessing the reliability of data, according to officials. GAO's AI Framework calls for management to provide reasonable assurance of the quality, reliability, and representativeness of the data used in the application, from its development through operation and maintenance. Addressing data sources and reliability is essential to model accuracy. Fully implementing the key practices can help DHS ensure accountable and responsible use of AI.

GAO-24-106246 Artificial Intelligence

Contents

Abbreviations

AI	artificial intelligence
AI Framework	GAO AI Accountability Framework
AIS	Automated Indicator Sharing
AS&F	Automated Scoring and Feedback
CIO	Chief Information Officer
CISA	Cybersecurity and Infrastructure Security Agency
CTOD	Chief Technology Officer Directorate
DHS	Department of Homeland Security
NDAA FY19	National Defense Authorization Act for Fiscal Year 2019
NDAA FY21	National Defense Authorization Act for Fiscal Year 2021
OMB	Office of Management and Budget
PIA	Privacy Impact Assessment
PII	Personally identifiable information

February 7, 2024

The Honorable Gary C. Peters
Chairman
The Honorable Rand Paul, M.D.
Ranking Member
Committee on Homeland Security and Governmental Affairs
United States Senate

The use of artificial intelligence (AI) in cybersecurity offers the potential to strengthen the nation's resilience against a range of cyber threats. For example, AI applications may be used to detect and defend against cyberattacks. At the same time, they may face threats from such attacks. While responsible AI use has the potential to help solve urgent challenges and make the world more prosperous, productive, innovative, and secure, irresponsible use could exacerbate societal harms, displace and disempower workers, stifle competition, and pose risks to national security.[1] We previously reported that cyber threats can arise from malicious actors seeking financial, political, or military gain.[2] Our report noted that adversaries would be highly motivated to exploit cyber defense systems that are based on machine learning algorithms. Federal agencies' efforts will be crucial in developing software that can help identify and address vulnerabilities while detecting and defending against attacks.

The Department of Homeland Security (DHS) is the lead federal agency responsible for defending against cyberattacks on critical U.S. infrastructure. DHS's Cybersecurity and Infrastructure Security Agency (CISA) is responsible for operations and coordination across federal entities to ensure the security and resilience of the nation's critical infrastructure. The mission of the Cybersecurity Division within CISA is to defend and secure cyberspace by leading national efforts to drive and enable effective national cyber defense, resilience of national critical

---

[1]Exec. Order No. 14110, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, § 1, (Oct. 30, 2023), 88 Fed. Reg. 75,191 (Nov. 1, 2023).

[2]GAO, *Technology Assessment: Artificial Intelligence: Emerging Opportunities, Challenges, and Implications,* GAO-18-142SP (Washington, D.C.: Mar. 28, 2018). See appendix II for a snapshot on the benefits and challenges with the use of AI in cybersecurity.

functions, and a robust technology ecosystem.[3] According to DHS, CISA operates multiple AI systems at various stages of development, including systems related to cybersecurity.[4]

The landscape of AI in the federal government is evolving rapidly. In August 2023, DHS issued a policy statement[5] to direct actions for DHS components to establish policy and practices governing the acquisition and use of AI within DHS.[6] In October 2023, the White House issued Executive Order No. 14110 on the *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, which, among other things, called on federal agencies to lead both the advancement of AI development and efforts to mitigate risks related to its development and use. It also sets new policies and principles for the responsible development and use of AI.[7]

Assessing agencies' current efforts to implement AI is a first step towards ensuring the federal government successfully leverages AI to accomplish its goals. GAO developed the AI Accountability Framework (the AI Framework) to help managers ensure accountability and the responsible use of AI in government programs and processes.[8] Organized around

[3]Cybersecurity and Infrastructure Security Agency (CISA), "About CISA," accessed August 8, 2023, https://www.cisa.gov/about.

[4]See DHS, *Artificial Intelligence Use Case Inventory,* accessed on November 14, 2023, https://www.dhs.gov/data/AI_inventory.

[5]DHS Policy Statement 139-06, *Acquisition and Use of Artificial Intelligence and Machine Learning by DHS Components,* (Aug. 8, 2023).

[6]For the purposes of this report, we use the term *component* in two different ways. We use the term to refer to components (offices or subdivisions) within the Department of Homeland Security, of which CISA is one. We also use the term to mean a part of an AI system.

[7]Exec. Order No. 14110, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, (Oct. 30, 2023), 88 Fed. Reg. 75,191 (Nov. 1, 2023).

[8]GAO, *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities,* GAO-21-519SP (Washington, D.C.: June 30, 2021). We developed the AI framework based on the following sources: (1) literature on accountability, governance frameworks, and principles on the use of AI; (2) presentations by and comments made by forum experts during a Comptroller General's forum; (3) interviews with subject matter experts including federal auditors and program managers, a state auditor, civil liberties advocates, industry representatives and legal counsel, developers, privacy experts, and data scientists; (4) GAO auditing standards and federal internal controls; (5) technical review of the framework and an outline of the forum findings by forum participants, including officials from three federal agencies and two Offices of Inspectors General; and (6) internal review by GAO subject matter experts.

four complementary principles—governance, data, performance, and monitoring—the AI Framework emphasizes substantive approaches that those implementing AI, as well as auditors and third-party assessors, can take to ensure responsible and accountable use of AI systems.[9]

We performed this work under the authority of the Comptroller General to conduct a review of DHS's use of AI for cybersecurity. This report examines the extent to which DHS (1) verified the accuracy of its inventory of AI systems used for cybersecurity and (2) incorporated selected practices from the AI Framework to manage and oversee its use of AI for cybersecurity.

To address these objectives, we reviewed relevant laws, Executive Orders, Office of Management and Budget (OMB) guidance and memorandums, agency policies and documents; and interviewed relevant agency officials. For the first objective, we reviewed DHS's 2022 AI Use Case Inventory to identify and review cybersecurity-related AI systems. We then reviewed agency documents and spoke with relevant officials at the Chief Technology Officer Directorate (CTOD), the office responsible for reporting the AI Use Case Inventory, to understand DHS's process to verify its use cases.

For the second objective, we applied 11 selected practices from the AI Framework to CISA's AI component—Automated Personally Identifiable Information (PII) Detection. We selected the 11 key practices to reflect all four principles from the AI Framework, align with early stages of AI adoption, and be highly relevant to the specific use case. For each selected practice, we considered (1) pertinent criteria from the AI Framework, National Institute for Standards and Technology and OMB guidance, and AI executive orders, and (2) associated relevant key questions from the AI Framework. We then assessed the key questions to the Automated PII Detection use case to determine whether the practice was:

- fully implemented—the agency provided evidence which showed that it fully or largely addressed key considerations;

---

[9]For each principle, the AI framework includes the following: key practices, which we developed by synthesizing information and identifying at least two sources that noted the importance of a certain practice in implementing AI systems; key questions, which we developed from information provided during a Comptroller General's forum, interviews with experts, and documents; and audit procedures, which we developed by reviewing the types of evidence noted in the Government Auditing Standards.

- mostly implemented—the agency provided evidence that it had addressed most of the key considerations;

- partially implemented—the agency provided evidence that it had addressed at least some of the key considerations;

- minimally implemented—the agency provided evidence that it had addressed at least one of the key considerations;

- not implemented—the agency did not provide evidence that it had addressed any of the key considerations.[10]

In conducting this analysis, we reviewed documentation from CISA which included technical specifications and requirements, workflows, data characterization, and test plans for the AI use case, as well as documentation on strategic and implementation plans. We interviewed CISA officials responsible for managing and overseeing AI aspects related to each key practice. For more information about our objectives, scope, and methodology, see appendix I.

We conducted this performance audit from September 2022 to February 2024 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 objectives.

Background

Federal use of AI has grown in recent years to advance automation, enhance data analysis, and improve government services, according to the Congressional Research Service.[11] For example, the National Defense Authorization Act for Fiscal Year 2021 (NDAA FY21) included provisions addressing various defense- and security-related AI activities,

---

[10]For some practices, not all key questions to consider were assessed due to the nature of the use case.

[11]Congressional Research Service, *Artificial Intelligence: Background, Selected Issues, and Policy Considerations*, R46795 (May 19, 2021).

in addition to the expansive National Artificial Intelligence Initiative Act of 2020.[12]

The White House has also issued three Executive Orders which sought to establish a coordinated federal government AI strategy guided by specific principles and promote the use of trustworthy AI across the federal government. Specifically:

- Executive Order No. 13859, *Maintaining American Leadership in Artificial Intelligence*, (February 2019) directed federal agencies to promote sustained investment in AI research and development in collaboration with non-federal entities, enhance access to federal data and computing resources, reduce barriers to the use of AI technologies, ensure that technical standards minimize vulnerabilities, train the next generation of American AI researchers, and develop action plans to protect American advantages in critical AI technology development.[13]

- Executive Order No. 13960, *Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government*, (December 2020) established a common policy for implementing principles related to lawfulness, performance, accuracy, reliability, safety, resilience, understandability, responsibility, transparency, accountability, and monitoring. In addition, it calls on the General Services Administration and Office of Personnel Management to expand AI expertise at agencies across government.[14] It also required agencies to create publicly available inventories of all non-classified and non-sensitive uses of AI (also known as use cases) in accordance with guidance

---

[12]National Artificial Intelligence Initiative Act of 2020, Division E of the William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021, Pub. L. No. 116-283, Div. E, § 5101(a), 134 Stat. 3388, 4524 (2021) codified at 15 U.S.C. § 9411. The National Artificial Intelligence Initiative Act of 2020 was enacted as Division E of the William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021. Although there have been numerous other recent laws and Executive Orders addressing the use of AI in the U.S., the purposes of this act made clear that its purpose was a comprehensive national strategy led by the National Artificial Intelligence Initiative Office. 15 U.S.C. §§ 9411(a) and 9412.

[13]Exec. Order No. 13589, *Maintaining American Leadership in Artificial Intelligence*, (Feb. 11, 2019), 84 Fed. Reg. 3,967 (Feb. 14, 2019).

[14]Exec. Order No. 13960, *Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government*, (Dec. 3, 2020), 85 Fed. Reg. 78,939 (Dec. 8, 2020), identified the responsible agencies as those defined in 44 U.S.C. § 3502 (1), but excludes those identified by 44 U.S.C. § 3502 (5) as independent regulatory agencies.

developed by the Federal Chief Information Officers Council.[15] Agencies are directed to prepare, identify, review, assess, and share inventories to the extent practicable and consistent with applicable law and policy, including those concerning the protection of privacy and of sensitive law enforcement, national security, and other protected information. The Executive Order also states that the principles for AI use for purposes other than national security and defense are to ensure that federal AI uses are consistent with our nation's values and beneficial to the public.[16]

- Executive Order No. 14110, *Safe, Secure, and Trustworthy Development and Use of AI*, (October 2023) established eight guiding principles and priorities for federal agencies to adhere to in carrying out activities using AI. These principles and priorities include: safety and security; promoting innovation and competition; supporting workers; advancing equity and civil rights; protecting consumers, patients, passengers, and students; protecting privacy and civil liberties; advancing federal government use of AI; and strengthening American leadership abroad. The Executive Order also contained specific measures related to cybersecurity. For example, it required relevant agencies to coordinate with the Director of CISA to assess potential risks related to the use of AI in critical infrastructure systems, including the ways in which deploying AI may make these systems more vulnerable to cyberattacks.[17]

In accordance with Executive Order 13960, DHS published its first AI Use Case Inventory in 2022. DHS components reported a total of 21 AI use cases, two of which were cybersecurity-related (Automated Scoring and Feedback, and Automated PII Detection).[18]

---

[15]Exec. Order No. 13960, § 5(b), 85 Fed. Reg. at 78,941.

[16]Exec. Order No. 13960, §§ 5(e), and 1, 85 Fed. Reg. at 78,941 and 78,939.

[17]Exec. Order No. 14110, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, (Oct. 30, 2023), 88 Fed. Reg. 75,191 (Nov. 1, 2023). GAO did not assess Exec. Order No. 14110 for this report because it was issued late in our review.

[18]According to DHS, Automated Scoring and Feedback and Automated PII Detection were initially developed in 2015. The latest versions of both use cases had been operating for fewer than 6 months at the time DHS developed the 2022 AI Use Case Inventory. The discussion of Automated Scoring and Feedback and Automated PII Detection will be described further later in this report.

DHS's AI Accountability Efforts

In 2020, DHS developed an AI strategy highlighting the following goals:

- assess the potential impact of AI on the DHS enterprise;

- invest in AI capabilities;

- mitigate AI risks to the agency and to the nation;

- develop an AI workforce; and

- improve public trust and engagement.[19]

The 2020 strategy states that DHS is currently deploying and operating various AI systems and that DHS components should have measures in place to increase transparency, increase accountability, and regularly monitor AI systems for potential bias and error. Further, it states that these efforts should be a part of a consistent approach to ensure effective governance rather than ad hoc component practices.

Additionally, DHS announced in August 2023 a department-wide policy establishing principles for the responsible use of AI.[20] In the policy statement, DHS directed each component to identify a senior career employee or servicemember with appropriate technical expertise to participate in a newly established agency AI Policy Working Group. The identified officials are responsible for providing an updated inventory of current use cases and an accounting of all planned AI use cases within each respective component. The policy states that the working group will assess the need for DHS components to update or revise their existing policies, procedures, and processes for the responsible acquisition and use of AI technologies.

---

[19]Department of Homeland Security (DHS), *Artificial Intelligence Strategy,* 2 (Dec. 3, 2020).

[20]DHS Policy Statement 139-06 states that DHS systems, programs, and activities using AI will conform to the requirements of Exec. Order No. 13960 and will only acquire and use AI in a manner that is consistent with the U.S. Constitution and all other applicable laws and policies. The policy also states that DHS will not collect, use, or disseminate data used in AI activities, or establish AI-enabled systems that make or support decisions, based on the inappropriate consideration of race, ethnicity, gender, national origin, religion, sexual orientation, gender identity, age, nationality, medical condition, or disability. DHS Policy Statement 139-06, *Acquisition and Use of Artificial Intelligence and Machine Learning by DHS Components,* § II at 2 (Aug. 8, 2023).

GAO's AI Framework

GAO published its AI Framework in June 2021 to help managers across the federal government ensure accountability and responsible use of AI in government programs and processes.[21] The AI Framework identifies key practices involved in the design, development, deployment, and continuous monitoring of AI systems.[22] Each practice includes a set of questions for entities, auditors, and third-party assessors to consider, along with audit procedures and types of evidence for auditors and third-party assessors to collect.[23]

The following 11 key practices were considered in DHS's management, operations, and oversight of its AI component for cybersecurity (see fig. 1).

---

[21]GAO-21-519SP.

[22]The term AI has a range of meanings in the scientific literature. The National Artificial Intelligence Initiative Act of 2020, Division E of the William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021 (NDAA FY21), Pub. L. No. 116-283, Div. E, § 5002(3), 134 Stat. 3388, 4524 (2021), defines AI as: a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to—(A) perceive real and virtual environments; (B) abstract such perceptions into models through analysis in an automated manner; and (C) use model inference to formulate options for information or action. 15 U.S.C. § 9401(3). An earlier definition of artificial intelligence, for the purposes of Department of Defense activities, was established in Section 238(g) of the John S. McCain National Defense Authorization Act for Fiscal Year 2019 (NDAA FY19), Pub. L. No. 115-232, § 238(g), 132 Stat. 1636,1697-98 (2018). GAO's AI Framework relies on a set of generalized characteristics of AI that are broader than the NDAA FY21 or NDAA FY19 enacted definitions; the AI Framework describes AI as having three distinct waves of development.

[23]The AI Framework distills insights from 23 cross-sectoral experts convened during the Forum on Artificial Intelligence by the Comptroller General of the United States held on September 9 and 10, 2020. To develop the AI Framework, we also conducted an extensive literature review and independent validation of key practices from program officials and subject matter experts.

Figure 1: Principles, Selected Key Practices, and Questions to Consider for Managing and Overseeing Artificial Intelligence

Principle	Practice	Key considerations
Governance	**Clear goals**	What goals and objectives does the entity expect to achieve throughout the AI life cycle? To what extent do stated goals and objectives represent a balanced set of priorities and adequately reflect stated values? How does the AI system help the entity meet its goals and objectives? To what extent does the entity communicate its AI strategic goals and objectives to the community of stakeholders? To what extent does the entity have the necessary resources to achieve the goals and objectives outlined for the AI life cycle? To what extent does the entity consistently measure progress towards stated goals and objectives?
	Roles and responsibilities	What are the roles, responsibilities, and delegation of authorities of personnel involved throughout the AI life cycle? To what extent has the entity clarified the roles, responsibilities, and delegated authorities to relevant stakeholders?
	Stakeholder involvement	What factors were considered when identifying the community of stakeholders involved throughout the life cycle? Which stakeholders did the entity include throughout the life cycle? What specific perspectives did stakeholders share, and how were they integrated throughout the life cycle? To what extent has the entity addressed stakeholder perspectives on the potential negative impacts of the AI system on end users and impacted populations?
	Technical specifications	What challenge/constraint is the AI system intended to solve? To what extent has the entity clearly defined technical specifications and requirements for the AI system? How do the technical specifications and requirements align with the AI system's goals and objectives? What justifications, if any, has the entity provided for the assumptions, boundaries, and limitations of the AI system?
	Compliance	To what extent has the entity identified the relevant laws, regulations, standards, and guidance, applicable to the AI system's use? How does the entity ensure that the AI system complies with relevant laws, regulations, standards, federal guidance, and policies? To what extent is the AI system in compliance with applicable laws, regulations, standards, federal guidance, and entity policies?
Data	**Sources**	How has the entity documented the AI system's data provenance, including sources, origins, transformations, augmentations, labels, dependencies, constraints, and metadata?
	Reliability	To what extent are data used to develop the AI system accurate, complete, and valid?
Performance	**Component-level documentation**	How is each model component solving a defined problem? How are the operating specifications and parameters of model and non-model components selected, evaluated, and optimized? How suitable are the components to the available data and operating conditions?
	System-level documentation	To what extent has the entity documented the AI system's development, testing methodology, metrics, and performance outcomes? To what extent does the documentation describe test results, limitations, and corrective actions, including efforts to minimize undesired effects in the outcomes?
	Human supervision	How has the entity considered an appropriate degree of human involvement in the automated decision-making processes? What procedures have been established for human supervision of the AI system? To what extent has the entity followed its procedures for human supervision to ensure accountability?
Monitoring	**Planning**	What plans has the entity developed to monitor the AI system? To what extent do the plans describe processes and procedures to continuously monitor the AI system? What is the established frequency for monitoring the AI system?

Source: GAO *AI Accountability Framework*; GAO (icons). | GAO-24-106246

Note: The selected key practices reflect all four principles in GAO's AI Accountability Framework, align with early stages of AI adoption and are relevant to the specific use case. For some practices, not all key questions to consider were assessed due to the nature of the use case.

DHS Has Not Taken Steps to Verify Whether Use Cases Are AI

According to Executive Order 13960, agencies including DHS are required to submit an inventory of their use cases of AI and publish these on the agency's website. We found the agency's inventory of AI uses cases for cybersecurity is not accurate. Although DHS has a review process for the inventory, the agency acknowledges that it does not confirm whether use cases identified by components are correctly characterized as AI as a part of this process.

CTOD Reviewed Submissions for the 2022 AI Use Case Inventory

DHS documentation shows the agency's CTOD reviewed information submitted by DHS components for the 2022 AI Use Case Inventory. According to DHS officials and documents, CTOD—which is responsible for compiling and reporting on the inventory—reviewed 79 AI use cases identified by DHS components and removed those that did not meet the criteria defined in Executive Order 13960 or the 2021 Chief Information Officers (CIO) Council's Guidance.[24] For example, CTOD removed use cases if a component flagged it as sensitive or if the use case was purely for research and development purposes.[25] Additionally, CTOD officials

---

[24]In 2020, the Office of Management and Budget (OMB) Chief Information Officers (CIO) Council issued guidance to federal agencies for creating their AI Use Case Inventories. The guidance provides the criteria and format for the inventory designated in Exec. Order No. 13960. The CIO Council issued updated guidance in 2023. OMB Federal Chief Information Officers Council, *2021 Guidance for Creating Agency Inventories of Artificial Intelligence Use Cases* (Oct. 6, 2020) (2021 CIO Council Guidance), and OMB Federal Chief Information Officers Council, *Guidance for Creating Agency Inventories of AI Use Cases Per EO 13960* (Aug. 2023) (2023 CIO Council Guidance). In previous work on federal agency implementation of the federal law and guidance related to AI, including selected requirements in Exec. Order No. 13960, GAO found that as of December 2023, DHS had implemented some but not all requirements related to Exec. Order No. 13960. Specifically, DHS prepared its AI Use Case Inventory and made it publicly accessible but did not include each of the required data elements for the AI Inventory and had not developed a plan to align with Exec. Order No. 13960 or retire AI applications found to be inconsistent with the Executive Order. GAO, *Artificial Intelligence: Agencies Have Begun Implementation but Need to Complete Key Requirements,* GAO-24-105980 (Washington, D.C.: December 12, 2023).

[25]According to Exec. Order No. 13960, agencies shall inventory only non-classified and non-sensitive use cases. Inventories shall not include AI research and development activities.

reviewed the descriptions of the AI use cases submitted by components with the National Defense Authorization Act for Fiscal Year 2019 (NDAA FY19) definition of AI, and spoke with components' subject matter experts to gather more information on their submissions.[26] After the CTOD's review, the final 2022 inventory contained 21 AI use cases and was published on the DHS website.[27]

In the final 2022 Use Case Inventory, DHS included two cybersecurity-related AI components—AIS Automated Scoring and Feedback (AS&F), and Automated PII Detection (see text box). Both are part of a broader system known as Automated Indicator Sharing (AIS). AIS, operated by CISA, shares cyber threat information in real time with the public and private sectors, information-sharing and analysis centers, and foreign government partners and companies. According to CISA, as more information is shared, participants become better informed and prepared to prevent damage related to cyber incidents. We initially selected both cybersecurity-related AI uses for our review, based on the 2022 AI Use Case Inventory.

DHS 2022 Artificial Intelligence Use Cases for Cybersecurity

Automated Scoring and Feedback (AS&F) is a predictive model that enriches information submitted to the Automated Indicator Sharing (AIS) system, which is designed to share cyber threat information. According to DHS, AS&F enriches information (1) by assessing whether the information can be corroborated with other sources available to the entity submitting the information and (2) by providing a confidence score that states the submitter's confidence in the accuracy of information they submit into AIS.

Automated PII Detection processes language to recognize personally identifiable information (PII) in cyber threat indicator submissions. A component of AIS, Automated PII Detection flags possible PII for an analyst to review. Analysts can then confirm or deny whether the system properly identified PII and remediate the information if the PII is not directly related to a cyber threat.

Source: GAO analysis of DHS documents. | GAO-24-106246

---

[26]Officials developed an information collection tool and sent it to components, along with instructions on how to populate the tool with data for the AI Use Case Inventory.

[27]Our review focuses on the 2022 AI Use Case Inventory. In 2023, DHS updated its Use Case Inventory, which as of November 2023 includes 50 AI use cases. The current AI Use Case Inventory can be accessed and downloaded at https://www.dhs.gov/publication/ai-use-case-inventory.

CTOD Did Not Verify the Accuracy of Components' Submissions for Inclusion in the 2022 AI Use Case Inventory

Although CTOD had a process to review components' submissions to the AI Use Case Inventory, we found this process does not include steps to verify whether each Use Case Inventory submission was characteristic of AI. CTOD officials told us that they relied on (1) the components to ensure that their AI Use Case Inventory submissions were accurate and (2) an internal framework for IT acquisitions for all systems, including those on the AI Use Case Inventory.[28]

In relying on components to ensure the accuracy of their AI Use Case submissions, CTOD sometimes asked components to revise or review submissions. According to DHS documents, in some instances when revisions were requested, some components did not reply or approve of additional changes made by CTOD. Further, we found the internal IT acquisitions framework that CTOD relied on to review the AI Use Case Inventory did not specifically address whether a system or component has the characteristics of AI.

Of the two cybersecurity-related AI systems included in the final 2022 AI Use Case Inventory, one did not have the characteristics of AI. As noted above, we initially included both AS&F and Automated PII Detection in our review. However, in July 2023, following a series of interviews and document requests, CISA staff said that AS&F was not developed or implemented as an AI component. Therefore, they did not consider it to be AI.

CISA Mission Engineering staff also stated that they did not understand the process for removing a system from the agency's AI Use Case Inventory. They told us they deferred to the CISA Office of the Technical Director and the CISA Office of the Chief Technology Officer to remove it from the agency's AI Use Case Inventory. According to CTOD officials, they had included AS&F in the AI Use Case Inventory because it met Executive Order 13960's broad definition of AI, which incorporates the

---

[28]According to CTOD, it also relies on existing review processes conducted through CTOD's Program Health Assessments, which consider seven criteria: risk management, performance risk, human capital, requirements and delivery, IT governance, contract and acquisition, and cybersecurity and privacy. These review and oversight efforts are informed by the governance of IT acquisition across the department, according to DHS.

definition from the NDAA FY19.[29] As of November 2023, DHS's publicly accessible AI Use Case Inventory still included AS&F as an AI use case.

According to OMB information quality guidelines—issued to ensure consistency with the Paperwork Reduction Act—and Memorandum M-19-15, agencies must embrace a basic standard of quality and consider quality in their information dissemination practices.[30] The law, the guidelines, and the memorandum explain that prudent decision-making depends on reliable, high-quality information. Agencies also must use quality assurance procedures before disseminating information publicly. Quality consists of utility (the data's utility for its intended users and for its intended purpose), integrity (security of the data), and objectivity (whether the disseminated information is accurate, reliable, and unbiased as a matter of presentation and substance).[31]

CTOD officials said they did not independently verify systems because they rely on components and existing IT governance and oversight efforts

---

[29]NDAA FY19, Pub. L. No. 115-232, § 238(g), 132 Stat. at 1697-98 (2018), 10 U.S.C. note prec. § 4061. Exec. Order No. 13960, incorporates by reference Section 238(g) of the NDAA FY19 to define AI. This definition includes the following: (1) any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets; (2) an artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action; (3) an artificial system designed to think or act like a human, including cognitive architectures and neural networks; (4) a set of techniques, including machine learning, that is designed to approximate a cognitive task; (5) an artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision-making and acting.

[30]Office of Management and Budget, OMB Memorandum M-19-15, *Improving Implementation of the Information Quality Act,* at 2 (Apr. 24, 2019), citing OMB, *Guidelines for Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies* (Guidelines), 67 Fed. Reg. 8452, 8458 (Feb. 22, 2002). Congress required OMB to issue guidelines consistent with the Paperwork Reduction Act, that provide policy and procedural guidance to agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by Federal agencies. Treasury and General Government Appropriations Act, 2001, Pub. L. No. 106-554, § 515(a), 114 Stat. 2763, 2763A-153 (2000) (as codified at 55 U.S.C. § 3516 note), citing to the purposes and provisions of the Paperwork Reduction Act, 44 U.S.C. § 3501(7) and (9).

[31]Pub. L. No. 106-554, § 515(a), 1114 Stat. at 2763A-154; 67 Fed. Reg. at 8453-54 and 8459-60; and OMB Memorandum M-19-15, at 2-3.

to ensure accuracy.[32] According to experts who participated in the Comptroller General's Forum on Artificial Intelligence, existing frameworks and standards may not provide sufficient detail on assessing social and ethical issues which may arise from the use of AI systems.[33] Further, DHS's August 2023 policy states that it will establish a working group to assess the need for components to update or revise their existing policies, procedures, and processes for the responsible, ethical, and authorized acquisition and use of AI and machine learning technologies across the DHS enterprise.[34]

Based on our review, the inclusion of AS&F raises questions about the overall reliability of DHS's AI Use Case Inventory.[35] Moreover, since DHS makes this information available on its public website, other agencies, third-party assessors, and the public also lack accurate information on the federal use of AI. Until it expands its process to include determining whether uses are correctly characterized as AI, DHS will be unable to ensure accurate reporting on its AI Use Case Inventory.

The one remaining cybersecurity system from the 2022 AI Use Case Inventory that DHS officials agreed was correctly characterized as AI was Automated PII Detection. In the section below, we apply the AI Framework's practices to this specific AI component.[36]

---

[32]These officials explained that they verified components' submissions against the criteria in relevant law and policy but did not confirm whether those submissions were characteristic of AI. Instead, they relied on the component's determination. The officials did not explain from which relevant laws or policy the criteria were drawn.

[33]GAO-21-519SP, 73-74. See appendix II on information shared by panel experts.

[34]DHS Policy Statement 139-06, *Acquisition and Use of Artificial Intelligence and Machine Learning by DHS Components*, § II at 4 (Aug. 8, 2023).

[35]A recent GAO report on federal implementation of key requirements in law and federal guidance, including Exec. Order No. 13960 described additional inaccuracies, such as missing information in DHS's AI Use Case Inventory. GAO made three recommendations to DHS on implementing federal guidance on the use of AI. DHS agreed with these recommendations. For additional information, see GAO-24-105980.

[36]As noted above, Automated PII Detection is an AI component embedded in the AIS system. For the purposes of this report, we refer to Automated PII Detection to include activities which may have been conducted for the AIS service, which CISA provides to enable real-time exchange of machine-readable cyber threat indicators and defensive measures between public and private-sector organizations.

CISA Applied Some but Not All Key Framework Practices to Oversee Its Use of AI for Cybersecurity

Based on our review of agency documents and interviews with CISA officials, we found that CISA fully implemented four of 11 key practices from GAO's AI Accountability Framework; implemented selected elements of five practices in the areas of governance, performance, and monitoring; and did not implement the data practices on sources and reliability of its Automated PII Detection component.[37] Figure 2 summarizes our assessment of CISA's implementation of selected practices from the AI Framework.

---

[37]In this report, we apply the AI Framework practices only to CISA's Automated PII component. As noted above, CISA officials told us they did not consider the other cyber component on its 2022 AI Use Case Inventory—AS&F—to be AI.

Figure 2: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Practices to Manage and Oversee Artificial Intelligence

PRINCIPLE	KEY PRACTICE	IMPLEMENTATION STATUS
Governance	Organizational	
	Clear goals	Mostly implemented
	Roles and responsibilities	Partially implemented
	Stakeholder involvement	Fully implemented
	System	
	Specifications	Fully implemented
	Compliance	Fully implemented
Data	Model development	
	Sources	Not implemented
	Reliability	Not implemented
Performance	Component level	
	Documentation	Mostly implemented
	System level	
	Documentation	Minimally implemented
	Human supervision	Fully implemented
Monitoring	Continuous monitoring	
	Planning	Partially implemented

Legend: Fully implemented, Mostly implemented, Partially implemented, Minimally implemented, Not implemented

Source: GAO analysis of agency documents and interviews with Department of Homeland Security officials; GAO (icons). | GAO-24-106246

Accessible text for Figure 2: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Practices to Manage and Oversee Artificial Intelligence

- Governance
  - Clear goals – Mostly implemented
  - Roles and responsibilities – Partially implemented
  - Stakeholder involvement – Fully implemented
  - Specifications – Fully implemented
  - Compliance – Fully implemented

- Data
  - Sources – Not implemented
  - Reliability – Not implemented
- Performance
  - Documentation (component) – Mostly implemented
  - Documentation (system) – Partially implemented
  - Human supervision – Fully implemented
- Monitoring
  - Planning – Partially implemented

Source: GAO analysis of agency documents and interviews with DHS officials. | GAO-24-106246

CISA Partially Implemented Selected AI Governance Practices Identified by the AI Framework

GAO's AI Framework calls for management and those charged with oversight of AI to promote accountability by establishing processes to manage, operate, and oversee implementation.[38] Governance helps entities ensure oversight and accountability, manage risks of AI, and ensure it meets performance requirements.[39]

Of the AI Framework's five selected governance practices, CISA implemented three: stakeholder involvement, system specifications, and compliance with relevant laws and guidance. It mostly implemented the practice of defining clear goals and partially implemented the practice of defining clear roles and responsibilities. Figure 3 summarizes our assessment of CISA's implementation of selected AI governance practices from the AI Framework, described below.

---

[38]GAO-21-519SP, 26.

[39]GAO-21-519SP, 26.

Figure 3: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Governance Practices

PRINCIPLE	KEY PRACTICE	IMPLEMENTATION STATUS
Governance — To help entities promote accountability and responsible use of AI systems	**Organizational**	
	Clear goals	Mostly implemented
	Roles and responsibilities	Partially implemented
	Stakeholder involvement	Fully implemented
	System	
	Specifications	Fully implemented
	Compliance	Fully implemented

Fully implemented   Mostly implemented   Partially implemented   Minimally implemented   Not implemented
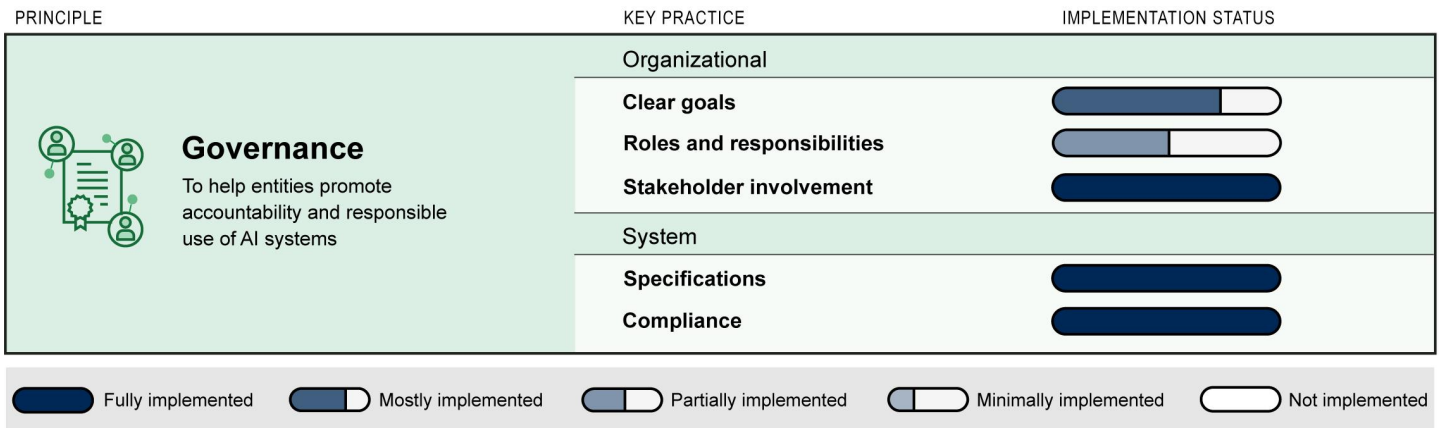
Source: GAO analysis of agency documents and interviews with Department of Homeland Security officials; GAO (icon). | GAO-24-106246

Accessible text for Figure 3: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Governance Practices

- Governance
  - Organizational
    - Clear goals – Mostly implemented
    - Roles and responsibilities – Partially implemented
    - Stakeholder involvement – Fully implemented
  - System
    - Specifications – Fully implemented
    - Compliance – Fully implemented

Source: GAO analysis of agency documents and interviews with DHS officials. | GAO-24-106246

Clear Goals

Key Considerations

- What goals and objectives does the entity expect to achieve by designing, developing, and/or deploying the AI system?
- To what extent do stated goals and objectives represent a balanced set of priorities and adequately reflect stated values?
- How does the AI system help the entity meet its goals and objectives?
- To what extent does the entity communicate its AI strategic goals and objectives to the community of stakeholders?
- To what extent does the entity have the necessary resources—funds, personnel, technologies, and time frames—to achieve the goals and objectives outlined for designing, developing and deploying the AI system?
- To what extent does the entity consistently measure progress towards stated goals and objectives?

Source: GAO AI Accountability Framework. | GAO-24-106246

According to the AI Framework, it is important for agencies to define clear goals and objectives for AI applications. Specifically, the AI Framework states that the goals and objectives should be specific and measurable to enable management to identify, analyze, and respond to risks related to achieving those objectives.[40] In addition, the entity should consider the resources necessary to achieve such goals.[41]

CISA defined clear goals and objectives for Automated PII Detection such as exchanging cyber threat indicators in a timely manner while protecting personal information. These goals and objectives represent a balanced set of priorities, such as timeliness and safety, based on criteria set forth

---

[40]GAO-21-519SP, 26-27.

[41]GAO-21-519SP, 32.

in the Cybersecurity Information Sharing Act of 2015.[42] According to CISA officials, Automated PII Detection automatically detects and flags potential PII that may be unrelated to the cyber threat, and is efficient and supports its goals.[43] CISA officials said that, without the Automated PII Detection component, screening for PII would take more time. CISA also communicated goals to external stakeholders by sharing content on their public website that addresses how AIS, including Automated PII detection, handles and processes PII.[44]

Further, CISA applied the agency's cost-estimating policies to determine resource needs (e.g., funds, personnel, and technology) for Automated PII Detection.[45] According to CISA officials, they took steps to review historical data and determine necessary resources for Automated PII

---

[42]The Cybersecurity Information Sharing Act of 2015 requires DHS to develop and maintain the capability to share real-time information on cyber threats. 6 U.S.C. § 1502. In response, DHS developed the AIS system, which receives and disseminates information on cyber threat indicators. The Cybersecurity Information Sharing Act of 2015 was enacted as Division N of the Consolidated Appropriations Act, 2016. Consolidated Appropriations Act, 2016, Pub. L. No. 114-113, 129 Stat. 2242 (2015), codified at 6 U.S.C. §§ 1500-1510.

[43]CISA analysts review flags and remove PII that is not directly related to a cyber threat.

[44]The Cybersecurity Information Sharing Act of 2015 required DHS to develop guidance for sharing cyber threat indicators among federal and non-federal entities. 6 U.S.C. § 1502. DHS developed such guidance and posted corresponding documents on its website, https://www.cisa.gov/resources-tools/resources/cybersecurity-information-sharing-act-2015-procedures-and-guidance. These documents include Non-Federal Entity Sharing Guidance; Privacy and Civil Liberties Final Guidelines; Federal Government Sharing Guidance; and Final Procedures Related to the Receipt of Cyber Threat Indicators and Defensive Measures by the Federal Government.

[45]The DHS National Protection and Programs Directorate requires program managers within DHS components to maintain accurate life cycle cost estimations for programs. These requirements include submitting an updated life cycle cost estimation to the National Protection and Programs Directorate annually. DHS Memorandum, Annual Life Cycle Cost Estimate Update, (Sept. 5, 2014).

Detection.[46] We reviewed CISA's cost-estimation results report that included these estimates for fiscal year 2023.[47]

However, CISA did not consistently measure progress toward its stated goals and objectives for timeliness. Although CISA's Office of Privacy assessed metrics related to the accuracy of PII screening—such as the frequency of cyber indicators correctly flagged and sent for human review, it did not develop metrics to assess goals related to timeliness. CISA officials told us that, consistent with the Cybersecurity Information Sharing Act of 2015, DHS continuously assesses the use of controls that may affect timeliness. According to DHS documents and officials, the objective of Automated PII Detection is to ensure cyber threat information is shared with AIS participants as quickly as possible while still protecting privacy. Officials also stated that Automated PII Detection is designed to facilitate the timely processing of cyber threat information by flagging PII. According to the AI Framework, consistently measuring progress toward stated goals and objectives can enable management to identify, analyze, and respond to risks related to achieving desired objectives.[48] Without metrics on all goals for Automated PII, CISA will be less likely to consistently measure progress and ensure intended outcomes are achieved.

---

[46]CISA officials told us that they determine necessary resources for the AI system, Automated PII Detection, within its broader AIS system.

[47]CISA develops a cost estimate for AIS within its National Cybersecurity Protection System (NCPS). CISA provided an NCPS Life Cycle Cost Estimate Update for Fiscal Year 2023. NCPS is an integrated program that delivers a range of capabilities, such as intrusion detection, analytics, information sharing, and intrusion prevention. These capabilities provide a technological foundation that enables CISA to secure and defend the Federal Civilian Executive Branch agencies' information technology infrastructure against advanced cyber threats. NCPS includes the hardware, software, supporting processes, training, and services that the program acquires, engineers, and supports to fulfill the agency's cybersecurity mission.

[48]GAO-21-519SP, 26-27, 32.

Roles and Responsibilities

Key Considerations

- What are the roles, responsibilities, and delegation of authorities of personnel involved in the design, development, deployment, assessment and monitoring of the AI system?
- To what extent has the entity clarified the roles, responsibilities, and delegated authorities to relevant stakeholders?

Source: GAO AI Accountability Framework. | GAO-24-106246

The AI Framework calls for agencies to define clear roles, responsibilities, and delegation of authority for the AI application.[49] Agencies are to do this at various stages of the AI life cycle, including design, development, deployment, assessment, and monitoring. The roles and responsibilities of personnel should be appropriate and clearly understood, according to the AI Framework and Executive Order No. 13960.[50]

We found CISA has not clearly defined the roles, responsibilities, and delegations of authority for all relevant personnel involved in managing and overseeing the implementation of the Automated PII Detection component. According to CISA's Mission Engineering subdivision—which is responsible for the technical design and development of the component—CISA uses a general project management process that defines roles and responsibilities for oversight of all systems within its subdivision. These officials stated that the project management process describes the general roles within their subdivision.[51] Based on our review, the project management process does not provide specific information on the roles, responsibilities, or delegation of authority for Automated PII Detection or AIS and does not define roles for other subdivisions. CISA officials told us there are other subdivisions involved in the management and oversight of Automated PII Detection, such as the CISA Office of Privacy. For these subdivisions, they provide briefings

---

[49]GAO-21-519SP, 26.

[50]Exec. Order No. 13960, Sec. 3(f) states that, "agencies shall ensure that human roles and responsibilities are clearly defined, understood, and appropriately assigned for the design, development, acquisition, and use of AI." See also the AI Framework at GAO-21-519SP, 27.

[51]CISA uses the Scale Agile Framework project management approach designed to accommodate continuous feedback and improvements to products. According to officials, this approach includes four installments throughout the year, and within those installments, sprint planning periods that reset every three weeks.

where they clarify roles and responsibilities associated with Automated PII.

According to the AI Framework, roles and responsibilities and delegations of authority should be clearly defined for all relevant stakeholders to ensure effective operations, timely corrections, and sustained oversight.[52] Without clearly defined roles, responsibilities, and delegation of authority for all the subdivisions involved in overseeing Automated PII Detection, CISA staff may find it difficult to ensure accountability over decision-making, implementation, and resolving issues, which may lead to unfavorable outcomes in using AI.

Stakeholder Involvement

Selected Key Considerations

- What factors were considered when identifying the community of stakeholders involved throughout the life cycle?
- Which stakeholders did the entity include throughout the design, development, deployment, assessment, and monitoring life cycle?
- What specific perspectives did stakeholders share, and how were they integrated across the design, development, deployment, assessment, and monitoring of the AI system?
- To what extent has the entity addressed stakeholder perspectives on the potential negative impacts of the AI system on end users and impacted populations?

Source: GAO AI Accountability Framework. | GAO-24-106246

According to the AI Framework, agencies are to include diverse perspectives from a community of stakeholders throughout the AI life cycle.[53] Strategies to incorporate diverse perspectives include establishing collaborative processes and multidisciplinary teams of subject matter experts in data science, software development, civil liberties, privacy and security, legal counsel, and risk management. These processes and multidisciplinary teams should also engage with individuals who may be using or operating the AI, or who may be affected by it.

---

[52]GAO-21-519SP, 26.

[53]GAO-21-519SP, 28.

CISA identified interagency stakeholders based on factors set forth in the Cybersecurity Information Sharing Act of 2015 and non-agency stakeholders based on their subject-matter relevance.[54] CISA engaged with interagency and non-agency stakeholders from 2015 to the present, to include engagement with a wide variety of stakeholders regarding recent changes to Automated PII Detection. Further, CISA consulted with interagency groups on matters related to privacy and civil liberties, prior to the more recent iteration of AIS in 2022. CISA initially engaged with stakeholders through various platforms, including public workshops, informal meetings, and comments via the *Federal Register* notice and comment process. CISA also engaged with internal stakeholders across other DHS components.

Further, CISA coordinated with stakeholders and considered perspectives related to potential negative outcomes of Automated PII Detection. For example, in 2016, CISA and the Department of Justice jointly published procedural documents that govern its handling of PII. CISA also consulted with private sector entities when developing its procedural documents for private sector participation in AIS between November 2015 and April 2016.

Technical Specifications

The AI Framework calls for agencies to include adequate documentation to define technical specifications.[55] It also calls for management to use judgment in determining the extent of documentation that is necessary to provide sufficient assurance that AI objectives will be met.[56]

---

[54]6 U.S.C. § 1501. Interagency stakeholders included the Departments of Justice, Defense, the Treasury, Commerce, and other federal entities. Non-agency stakeholders included private and non-governmental groups such as civil liberties organizations and technology and cybersecurity councils.

[55]GAO-21-519SP, 29.

[56]According to the AI Framework, entities define technical specifications to ensure the AI meets its intended purpose. GAO-21-519SP, 29.

Key Considerations

- What challenge/constraint is the AI system intended to solve?
- To what extent has the entity clearly defined technical specifications and requirements for the AI system?
- How do the technical specifications and requirements align with the AI system's goals and objectives?
- What justifications, if any, has the entity provided for the assumptions, boundaries, and limitations of the AI system?

Source: GAO AI Accountability Framework. | GAO-24-106246

CISA is using Automated PII Detection to address resource constraints due to the volume of information received by AIS, which creates constraints for timely human review, according to officials. CISA clearly defined and documented technical specifications for Automated PII Detection, and explained how these specifications support goals. For example, CISA documentation indicates how various technical specifications lessen the receipt of PII unnecessary to the cyber threat and enable automated screening for PII. Further, CISA officials noted that Automated PII Detection was designed to allow the dissemination of PII directly related to the cyber threat because sharing such information is permitted by the Cybersecurity Information Sharing Act of 2015 and could help AIS users understand the relevance and nature of a particular threat indicator. CISA officials explained that these specifications helped achieve Automated PII Detection goals, which included complying with the Cybersecurity Information Sharing Act of 2015 to ensure privacy protections and share real-time information on cyber threat indicators. Additionally, CISA officials provided justifications for limitations with Automated PII Detection such as those related to detecting and flagging potential PII within the cyber threat indicators.

Compliance

Key Considerations

- To what extent has the entity identified the relevant laws, regulations, standards, and guidance, applicable to the AI system's use?
- How does the entity ensure that the AI system complies with relevant laws, regulations, standards, federal guidance, and policies?
- To what extent is the AI system in compliance with applicable laws, regulations, standards, federal guidance, and entity policies?

Source: GAO AI Accountability Framework. | GAO-24-106246

According to the AI Framework, agencies are to ensure AI applications comply with relevant laws, regulations, standards, and guidance.[57] Agencies can take a proactive approach to ensuring compliance by considering applicable laws and regulations, industry standards, and guidance from federal agencies and other entities. Existing data privacy and non-discrimination laws are likely to be relevant for AI that processes personally identifiable information or sensitive data.[58]

CISA considered and documented compliance with relevant laws at a high level within its AIS system, which included considerations for PII detection capabilities.[59] According to DHS documentation and CISA officials, relevant laws include the Cybersecurity Information Sharing Act of 2015, the Privacy Act of 1974, the E-Government Act of 2002, and

---

[57]GAO-21-519SP, 30.

[58]For example, in October 2023, the Administration issued an Executive Order on the *Safe, Secure, and Trustworthy Development and Use of AI*, calling on federal agencies to lead both the advancement of AI development and efforts to mitigate risks related to its development and use. It also sets new policies and principles for the responsible use and development of AI in multiple categories such as: safety and security; promoting innovation and competition; supporting workers; advancing equity and civil rights; protecting consumers, patients, passengers, and students; protecting privacy and civil liberties; advancing federal government use; and strengthening American leadership abroad in AI-related use. Exec. Order No. 14110, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (Oct. 30, 2023), 88 Fed. Reg. 75,191 (Nov. 1, 2023).

[59]When changes are made to AIS, CISA's Privacy Office and other appropriate offices conduct compliance reviews, according to officials.

others.[60] DHS took steps to meet legal requirements of the Cybersecurity Information Sharing Act of 2015 by conducting biennial privacy reviews and the E-Government Act of 2002 by developing a Privacy Impact Assessment (PIA).[61] These steps are consistent with the Privacy Act of 1974, which establishes requirements regarding the collection, maintenance, use, and dissemination of personal information about individuals that is maintained in systems of records by federal agencies.

Further, CISA took steps to ensure ongoing compliance for Automated PII Detection. According to DHS's internal policy, CISA is to designate a

---

[60]The Cybersecurity Information Sharing Act of 2015 mandated that DHS develop procedures and maintain a capability for sharing of cyber threat indicators and defensive measures between the federal government and non-federal entities including the private sector. 6 U.S.C. § 1502(a). The Privacy Act of 1974, Pub. L. No. 93-579, 88 Stat. 1896, codified at 5 U.S.C. § 552a, is the principal law governing the handling of personally identifiable information (PII) contained within systems of records by federal agencies. The Privacy Act generally establishes requirements for all federal agencies regarding the collection, maintenance, use, and dissemination of PII. The Privacy Act pertains to AIS due to the risk of collecting unnecessary PII through AIS submissions. The Homeland Security Act of 2002, as amended, is DHS and CISA's authorizing statute. It establishes CISA, including the roles of the CISA Director, Executive Assistant Director for Cybersecurity, and CISA Office of Privacy, and authorizes the CISA Director and Executive Assistant Director for Cybersecurity to undertake various cybersecurity activities, including those related to information sharing. Homeland Security Act of 2002, as amended, Pub. L. No. 107-296, 116 Stat. 2135, codified in relevant part at 6 U.S.C. §§ 222, 652, and 659. The E-Government Act of 2002 also includes a relevant provision, Pub. L. No. 107-347, § 208, 116 Stat. 2899, 2921-22, 44 U.S.C. § 3501 note.

[61]Privacy Impact Assessments (PIA) are required by Section 208 of the E-Government Act for all federal government agencies that develop or procure new information technology involving the collection, maintenance, or dissemination of information in identifiable form or that make substantial changes to existing information technology that manages information in identifiable form. A PIA is an analysis of how information in identifiable form is collected, stored, protected, shared, and managed. The purpose of a PIA is to demonstrate that system owners and developers have incorporated privacy protections throughout the entire life cycle of a system. The act requires an agency to make PIAs publicly available, except when an agency in its discretion determines publication of the PIA would raise security concerns, reveal classified (i.e., national security) information, or sensitive (e.g., potentially damaging to national interest, law enforcement effort or competitive business interest contained in the assessment) information. E-Government Act of 2002, Pub. L. No. 107-347, § 208, 116 Stat. 2899, 2921-22, 44 U.S.C. § 3501 note. Section 222 of the Homeland Security Act of 2002, as amended, further requires the Chief Privacy Officer of the DHS to ensure that the technology used by DHS sustains privacy protections. 6 U.S.C. § 142. Agency responsibilities for completing the biennial compliance reports, including a quantitative and qualitative assessment of the effect of the sharing of cyber threat indicators or defensive measures with the federal government on privacy and civil liberties of specific individuals, including the number of notices that were issued with respect to a failure to remove information not directly related to a cybersecurity threat that was personal information of a specific individual or information that identified a specific individual are found at 6 U.S.C. § 1506(b)(2)(D)(iv).

Change Control Board—which includes officials in departments (e.g., the Departments of Justice and Defense)—to review certain changes made within the AIS system on a continual basis to identify and mitigate any privacy, civil liberties, and compliance concerns.[62] According to officials, the agency conducted such reviews in 2016 and prior to launching a more recent version of AIS in 2022. In addition, CISA officials said all changes to AIS undergo internal review processes to address such concerns, such as regular program briefings to the CISA Office of Privacy. According to the AI Framework, ensuring ongoing compliance helps entities demonstrate a commitment to principles and values that foster public trust in responsible AI use.[63]

CISA Did Not Implement Selected Practices for Appropriateness or Reliability of Data

GAO's AI Framework calls for management, and those charged with oversight of AI, to provide reasonable assurance of the quality, reliability, and representativeness of the data included in the application, from its development stage to its operation.[64] CISA did not implement either of the two selected data practices: documenting sources and ensuring reliability. Figure 4 summarizes our assessment of CISA's implementation of selected data practices from the AI Framework, described below.
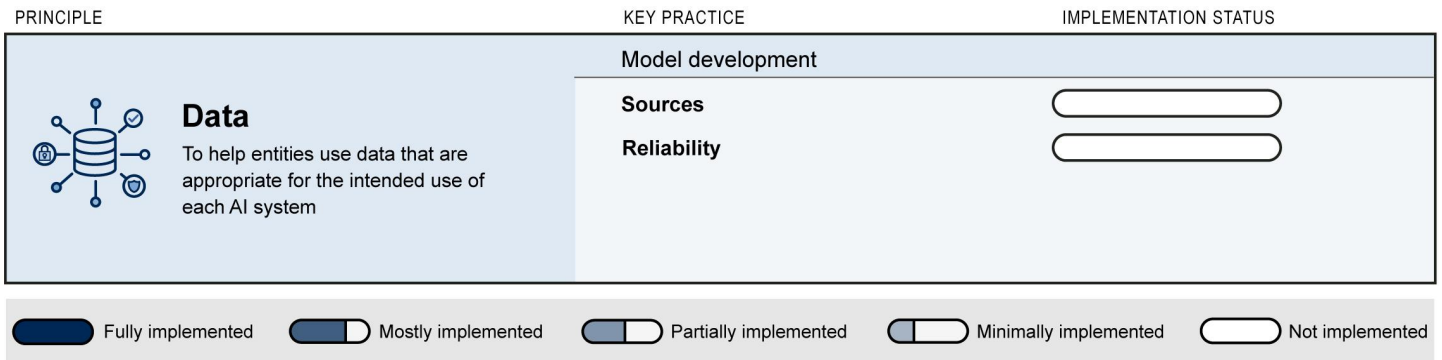
[62]The Department of Homeland Security and the Department of Justice, *Final Procedures Related to the Receipt of Cyber Threat Indicators and Defensive Measures by the Federal Government* (June 15, 2016), 4.

[63]GAO-21-519SP, 30.

[64]GAO-21-519SP, 38.

Figure 4: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Data Practices

PRINCIPLE	KEY PRACTICE	IMPLEMENTATION STATUS
Data To help entities use data that are appropriate for the intended use of each AI system	**Model development**	
	Sources	
	Reliability	

Legend: Fully implemented · Mostly implemented · Partially implemented · Minimally implemented · Not implemented

Source: GAO analysis of agency documents and interviews with Department of Homeland Security officials; GAO (icon). | GAO-24-106246

Accessible text of Figure 4: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Data Practices

- Principle
  - Data: To help entities use data that are appropriate for the intended use of each AI system
- Key Practice
  - Sources – Not implemented
  - Reliability – Not implemented

Source: GAO analysis of agency documents and interviews with DHS officials. | GAO-24-106246

Sources

Selected Key Considerations

- How has the entity documented the AI system's data provenance, including sources, origins, transformations, augmentations, labels, dependencies, constraints, and metadata?

Source: GAO AI Accountability Framework. | GAO-24-106246

According to the AI Framework, it is important for agencies and entities to document the sources and origins of data used to develop the models underpinning the AI.[65] In addition to documenting how the data were

---

[65]GAO-21-519SP, 39.

collected, entities should document how they were curated and used to increase transparency and accountability.

CISA did not document the sources and origins of data used to develop the PII detection capabilities. According to CISA officials, the data may have included information such as fabricated names, emails, and Social Security numbers to develop Automated PII Detection's ability to remove unnecessary PII. According to CISA officials, the personnel involved were no longer available and, due to its retention policy of maintaining records for 3 years, the officials no longer have access to the data. The initial development of Automated PII Detection was conducted more than 3 years ago, according to CISA.

According to the AI Framework, documenting the provenance and use of data in AI models can ensure data quality and enable third-party assessments.[66] Without documenting the source data, CISA cannot ensure outcomes are consistent and appropriate for detecting PII.

Reliability

> **Selected Key Considerations**
> - To what extent are data used to develop the AI system accurate, complete, and valid?
>
> Source: GAO AI Accountability Framework. | GAO-24-106246

The AI Framework states that agencies and entities should ensure that data used to develop AI models are reliable because data reliability affects the accuracy of model predictions. In addition, entities should implement procedures to reasonably ensure that data added to the system are complete, accurate, and valid.[67]

CISA officials told us they were not aware of any data reliability assessments conducted by staff or contractors and did not have any associated documentation. Nevertheless, these officials considered the data used in the development of Automated PII Detection to be adequate

---

[66]The term "data provenance" refers to a record that accounts for the origin of a piece of data (in a database, document, or repository), together with an explanation of how and why it got to the present place. A provenance record will document the history for each piece of data. GAO-21-519SP, 39.

[67]GAO-21-519SP, 40.

for its intended purpose of removing PII not directly related to the cyber threat, such as names, email addresses, and identification numbers.

According to the AI Framework, entities should assess the reliability of data used to develop the AI model because data reliability affects the accuracy of model predictions.[68] CISA officials told us they did not have additional information on data reliability efforts due to their retention policy of maintaining records for 3 years. Without such assessments, CISA limits its ability to ensure that the data are accurate, complete, and valid, thereby limiting its ability to mitigate unintended model outcomes.

CISA Did Not Fully Implement Selected Practices for System Performance Documentation

According to GAO's AI Framework, management and those charged with oversight of AI are to ensure results are consistent with program objectives.[69] Performance assessments can help agencies improve operations, reduce costs, facilitate decision-making by parties responsible for overseeing or initiating corrective action, and contribute to public accountability of AI.[70] Of the three selected performance practices, CISA mostly implemented the one on component-level documentation, minimally implemented the one on system-level documentation, and fully implemented the one on developing human supervision procedures. Figure 5 summarizes our assessment of CISA's implementation of selected practices for AI performance from the AI Framework, described below.
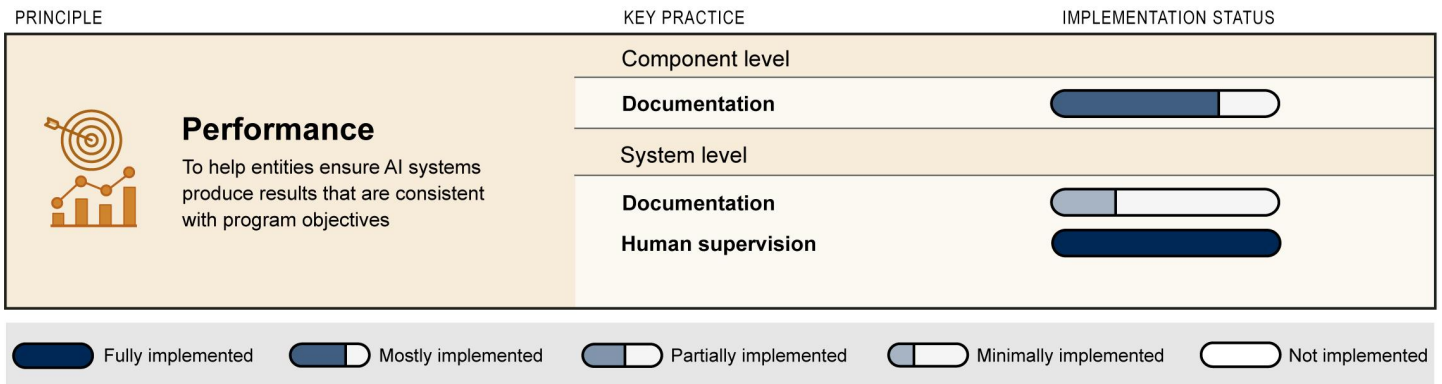
---

[68]GAO-21-519SP, 40.

[69]GAO-21-519SP, 48.

[70]GAO, *Government Auditing Standards 2018 Revision Technical Update April 2021*, GAO-21-368G (Washington, D.C.: Apr. 14, 2021).

Figure 5: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Performance Practices

Source: GAO analysis of agency documents and interviews with Department of Homeland Security officials; GAO (icon). | GAO-24-106246

Accessible text for Figure 5: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Performance Practices

- Performance: To help entities ensure AI systems produce results that are consistent with program objectives
- Key Practice
  - Component level
    - Documentation – Mostly implemented
  - System Level
    - Documentation – minimally implemented
    - Human supervision – Fully implemented

Source: GAO analysis of agency documents and interviews with DHS officials. | GAO-24-106246

Component Level Documentation

Selected Key Considerations

- How is each model component solving a defined problem?
- How are the operating specifications and parameters of model and non-model components selected, evaluated, and optimized?
- How suitable are the components to the available data and operating conditions?

Source: GAO AI Accountability Framework. | GAO-24-106246

The AI Framework states that agencies and entities should catalog the components of the AI and document the purpose of the components, including their specifications and parameters.[71]

CISA documentation showed that CISA considered how model elements addressed the problem of protecting PII in cyber threat indicators. CISA also selected and evaluated the operating specifications of the Automated PII Detection component. For example, the agency documented that it compared the speed and accuracy of several natural language processing solutions for identifying PII prior to selecting a solution.[72] Additionally, CISA considered whether the Automated PII Detection was appropriate and suitable for the operating conditions. Specifically, according to CISA documentation, all cybersecurity threat indicator submissions containing potential PII were initially reviewed manually by staff. But the use of manual checking could not be sustained due to the volume of data. In response, CISA adjusted the Automated PII Detection rules so that it would only send indicators for human review that included multiple types of PII.

However, CISA did not document whether specifications and parameters are appropriately optimized. Although the agency provided us with screenshots showing how changes to certain specifications, such as increased system memory, reduced the system's processing time, we could not verify that the screenshots had resulted from CISA's efforts to optimize processing time for the Automated PII Detection component. A senior official in the CTOD office said that the department followed existing acquisition processes that did not require this level of documentation and noted that Automated PII Detection was developed prior to the development of federal guidance on AI. However, we found that Automated PII Detection has undergone changes and updates as recently as May 2022. According to DHS's own Artificial Intelligence Strategy, components operating AI must continually validate the

---

[71]Performance assessment at the component level determines whether each component meets its defined objective. Components are technology assets that represent building blocks of an AI system and include hardware and software that apply mathematical algorithms to data. Performance assessment at the system level determines whether the components work well as an integrated whole. GAO-21-519SP, 48.

[72]Natural language processing is an AI technique that can detect words and meaning from text.

performance, monitor, and take action to mitigate risks posed by bias or other unintended outcomes.[73]

According to the AI Framework, documentation on model and non-model elements of the AI, along with specifications and parameters, provides assurance of the appropriateness of the components selected, enhances transparency, and increases users' and public trust in the AI system.[74] Without such documentation, CISA lacks sufficient assurance that the component is performing as it is intended to perform.

System Level Documentation

Key Considerations

- To what extent has the entity documented the AI system's development, testing methodology, metrics, and performance outcomes?

- To what extent does the documentation describe test results, limitations, and corrective actions, including efforts to minimize undesired effects in the outcomes?

Source: GAO AI Accountability Framework. | GAO-24-106246

According to the AI Framework, agencies and entities should document the methods for assessment, performance metrics, and outcomes of the AI application to provide transparency of its performance.[75] As AI components are integrated, entities should iteratively test the system as a whole and document the tests performed and the corresponding results.[76]

CISA documented the development of Automated PII Detection and some methods of assessing performance outcomes. Specifically, CISA documented how changes to the Automated PII Detection screening rules improved performance by reducing the number of findings that resulted in false positives.[77] But it did not document methods for testing performance prior to deployment. Although CISA provided us with five test reports and

---

[73]DHS, *Artificial Intelligence Strategy,* 9.

[74]GAO-21-519SP, 59.

[75]GAO-21-519SP, 7.

[76]GAO-21-519SP, 51, 52.

[77]The term "false positives" used here refers to instances of data being mistakenly identified as PII.

seven "tickets" documenting changes, most of the testing involved AIS (the larger cyber threat indicator system) rather than the PII detection. Two test reports were associated with Automated PII Detection. Of these, one documented testing of the Human Review functionality and not Automated PII Detection itself. The remaining test report documented the results of performance testing. But these tests were conducted after the most recent 2018 version of Automated PII Detection was implemented.[78] In addition, the documentation described the testing results for Automated PII Detection but did not document limitations or corrective actions taken to minimize undesired effects.

A senior official in the CTOD office said the department followed existing processes that did not require this level of documentation. However, according to DHS's Artificial Intelligence Strategy, components of operating AI systems must continually validate the performance of systems to monitor for and take actions to mitigate risks posed by bias or other unintended outcomes.[79]

Further, according to the AI Framework, documentation of performance testing is needed to ensure the AI performs reliably across a range of conditions and in a transparent manner.[80] Without such documentation on performance testing, CISA cannot ensure transparency over the performance of the component. In addition, third-party assessors and internal stakeholders may not be able to reproduce any assessments and testing results.

---

[78]The most recent version of Automated PII Detection deployed in 2018 is the version that flags indicators containing multiple types of PII for human review.

[79]DHS, *Artificial Intelligence Strategy,* 9.

[80]GAO-21-519SP, 51, 52.

<u>Human Supervision</u>

Key Considerations

- How has the entity considered an appropriate degree of human involvement in the automated decision-making processes?
- What procedures have been established for human supervision of the AI system?
- To what extent has the entity followed its procedures for human supervision to ensure accountability?

Source: GAO AI Accountability Framework. | GAO-24-106246

According to the AI Framework, agencies and entities should determine the appropriate degree of human supervision and establish procedures accordingly to ensure the system goals are met.[81] This degree depends on several factors, including the purpose and potential consequences of the AI.[82]

CISA considered the appropriate level of human involvement in Automated PII Detection. For example, CISA developed a set of standard operating procedures for human review. This documentation also defines procedures for human supervision of Automated PII Detection whereby CISA analysts review flagged cyber threat indicators identified by Automated PII Detection as having more than one type of PII in a single field of information. The analyst then removes any unnecessary PII before the indicator is disseminated. CISA documentation described procedures and policies for ensuring analysts adhere to privacy requirements and follow human supervision procedures.

CISA Partially Implemented the Selected Practice on Monitoring Planning

According to GAO's AI Framework, management, and those charged with oversight of AI, are to establish a monitoring framework to ensure the AI maintains its utility and remains aligned with current objectives.[83] CISA partially implemented the one selected monitoring practice on planning. Figure 6 summarizes our assessment of CISA's implementation of the
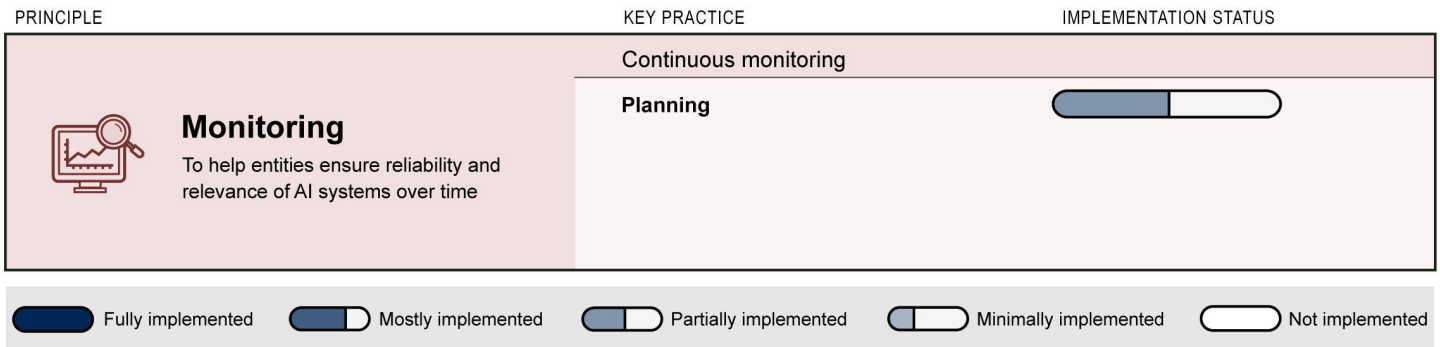
---

[81]GAO-21-519SP, 53.

[82]GAO-21-519SP, 53.

[83]GAO-21-519SP, 60.

selected practice for AI monitoring from the AI Framework, described below.

Figure 6: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Monitoring Practices

PRINCIPLE	KEY PRACTICE	IMPLEMENTATION STATUS
Monitoring To help entities ensure reliability and relevance of AI systems over time	Continuous monitoring	
	Planning	(partially implemented)

Fully implemented   Mostly implemented   Partially implemented   Minimally implemented   Not implemented

Source: GAO analysis of agency documents and interviews with Department of Homeland Security officials; GAO (icon). | GAO-24-106246

Accessible text for Figure 6: Status of the Cybersecurity and Infrastructure Security Agency's (CISA) Implementation of Selected Key Monitoring Practices

- Monitoring: To help entities ensure reliability and relevance of AI systems over time.
  - Key Practice
    - Continuous monitoring
      - Planning – Partially impemented

Source: GAO analysis of agency documents and interviews with DHS officials. | GAO-24-106246

Planning

> **Selected Key Considerations**
> - What plans has the entity developed to monitor the AI system?
> - To what extent do the plans describe processes and procedures to continuously monitor the AI system?
> - What is the established frequency for monitoring the AI system?
>
> Source: GAO AI Accountability Framework. | GAO-24-106246

The AI Framework states that agencies and entities should develop plans to monitor performance and risks continuously or routinely, including the

risk of bias and risks to privacy and security.[84] The plan should include a monitoring frequency that is appropriate for the use case.[85]

CISA took some steps to monitor Automated PII Detection. Specifically, CISA officials told us the agency monitors the performance of Automated PII Detection using a variety of methods across multiple components and at varying frequencies. For example, the performance of all production systems, including AIS, is monitored by the system operations team on a continual basis. Officials said the agency performs tests to ensure continued operation of the PII detection function using synthetic data. Additionally, officials pointed to the biannual Privacy Oversight Review of CISA's handling of PII in cybersecurity activities conducted by CISA's Office of Privacy as an example of monitoring that includes Automated PII Detection.

Although CISA documented a monitoring plan for AIS, the plan does not address monitoring of Automated PII Detection. CISA developed a monitoring plan for the AIS system to ensure it was operating as intended. However, the plan does not include specific procedures or frequencies for monitoring Automated PII Detection, which is a component of the larger AIS system. CISA officials stated that they do not develop specific plans or processes for smaller components like Automated PII Detection.

According to DHS's Artificial Intelligence Strategy, DHS components should take actions to continuously identify risks for new and existing systems.[86] For example, the strategy states DHS will ensure components have measures in place to regularly monitor AI systems for potential bias and error.[87]

In addition, according to the AI Framework, establishing frequencies to monitor the AI can help ensure that risks to privacy and security are identified and mitigated in a timely manner.[88] Without establishing such procedures to monitor the Automated PII Detection component in its AIS monitoring plan, CISA is limited in its ability to ensure that the component

---

[84]GAO-21-519SP, 61.

[85]GAO-21-519SP, 61.

[86]DHS, *Artificial Intelligence Strategy,* 2.

[87]DHS, *Artificial Intelligence Strategy,* 3.

[88]GAO-21-519SP, 61, 62.

remains useful, continues to align with current objectives, and addresses possible risks, such as bias, privacy, and security.

Conclusions

In its process of reviewing candidate use cases for inclusion in the department's inventory, DHS does not determine whether such cases are correctly characterized as AI. As a result, DHS cannot ensure the accuracy of the AI inventory.

In operating Automated PII Detection, DHS is fully implementing some key AI accountability practices but other important practices for governance, data, performance, and monitoring have not been fully implemented. The absence of critical data practices is particularly concerning because of the potential adverse effects on model results.

Ensuring responsible and accountable use of AI will be critical as DHS builds its capabilities to use AI for its operations. By fully implementing accountability practices, DHS can promote public trust and confidence that AI can be a highly effective tool for helping attain strategic outcomes.

Recommendations for Executive Action

We are making the following eight recommendations to DHS regarding its AI inventory and implementation of AI Framework practices on governance, data, performance, monitoring:

The Chief Technology Officer should expand its review process to include steps to verify the accuracy of its AI inventory submissions. (Recommendation 1)

Governance:

- The Director of CISA should develop metrics to consistently measure progress toward all stated goals and objectives for Automated PII Detection. (Recommendation 2)

- The Director of CISA should clearly define the roles and responsibilities and delegation of authority of all relevant stakeholders involved in managing and overseeing the implementation of the Automated PII Detection component to ensure effective operations and sustained oversight. (Recommendation 3)

Data:

- The Director of CISA should document the sources and origins of data used to develop the Automated PII Detection component. (Recommendation 4)

- The Director of CISA should take steps to assess and document the reliability of data used to enhance the representativeness, quality, and accuracy of the Automated PII Detection component. (Recommendation 5)

Performance:

- The Director of CISA should document its process for optimizing the elements used within the Automated PII Detection component. (Recommendation 6)

- The Director of CISA should document its methods for testing performance including limitations, and corrective actions taken to minimize undesired effects of the Automated PII Detection component to ensure transparency about the system's performance. (Recommendation 7)

Monitoring:

- The Director of CISA should establish specific procedures and frequencies to monitor the Automated PII Detection component to ensure it performs as intended. (Recommendation 8)

Agency Comments

We provided a draft of this report to the Department of Homeland Security for review and comment. In its comments reproduced in appendix III, DHS concurred with our recommendations. In addition, we received technical comments from DHS, which we incorporated into the draft, as appropriate. The following summarizes DHS's written responses to our recommendations:

- DHS concurs with Recommendation 1. DHS stated that CTOD updated the process for reviewing and adding use cases in Fall 2023 and implemented new oversight mechanisms to identify new AI use cases. As an example, DHS cited expansion of its central internal repository of use cases that provides interim updates to the public inventory. The repository now includes new additions that then undergo a new process for review and approval. In addition, CTOD

will further update the review process and evaluation criteria to ensure it is standardized and rigorous.

- DHS concurs with Recommendation 2 and stated that CISA's Cybersecurity Division (CSD) Mission Engineering (ME) Subdivision, in collaboration with CISA's Privacy Officer and other CISA stakeholders, will determine a method for documenting and tracking metrics to measure the goals and objectives for Automated PII Detection and will develop metrics to consistently measure the stated goals and objectives for Automated PII Detection.

- DHS concurs with Recommendation 3 and stated that CISA CSD ME will develop a document that defines the roles and responsibilities of the stakeholders involved in managing and overseeing the implementation of the Automated PII Detection component.

- DHS concurs with Recommendation 4 and stated that CISA CSD ME will document the sources and origins of data used to develop the Automated PII Detection component.

- DHS concurs with Recommendation 5 and stated that CISA CSD ME will assess and document the reliability of data used to enhance the representativeness, quality, and accuracy of the Automated PII Detection component.

- DHS concurs with Recommendation 6 and stated that CISA CSD ME will develop a document that defines the process for optimizing the elements used with the Automated PII Detection component.

- DHS concurs with Recommendation 7 and stated that CISA CSD ME will develop test planning documentation to further define the methods for testing performance of the Automate PII Detection component.

- DHS concurs with Recommendation 8 and stated that CISA CSD ME, in coordination with the CSD Technology Director, the CISA Privacy Officer, and other CISA stakeholders will update CISA's AIS monitoring plans to include procedures and frequencies specific to monitoring the Automated PII Detection component.

We are sending copies of this report to the appropriate congressional committees and the Secretary of Homeland Security. In addition, the report is available at no charge on the GAO website at https://www.gao.gov.

If you or your staff have any questions about this report, please contact me at (202) 512-6888 or WrightC@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on

the last page of this report. GAO staff who made key contributions to this report are listed in appendix IV.

Candice N. Wright
Director, Science, Technology Assessment, and Analytics

Kevin Walsh
Director, Information Technology and Cybersecurity

Appendix I: Objectives, Scope, and Methodology

This report examines the extent to which the Department of Homeland Security (DHS) (1) verified the accuracy of its inventory of Artificial Intelligence (AI) systems used for cybersecurity and (2) the extent to which DHS incorporated selected practices from GAO's AI Accountability Framework for Federal Agencies and Other Entities (the AI Framework) to manage and oversee its use of AI for cybersecurity.

For the first objective, we reviewed DHS's 2022 AI Use Case Inventory and related documentation to identify cyber-security related AI use cases and review the process used for developing the inventory. Documents we reviewed included DHS guidance and an information collection tool populated with information used to compile the inventory. We also interviewed agency officials to assess DHS's verification process for its AI Use Case Inventory. During interviews with relevant Chief Technology Officer Directorate (CTOD) and CISA officials, we asked questions about the process by which agency components submit use cases to the inventory, and how the agency verifies these submissions. We also reviewed Executive Order Nos. 13859, and 13960, relevant Office and Management and Budget (OMB) guidance and memorandums.

For the second objective, we applied 11 practices from GAO's AI Framework to CISA's AI component—Automated PII Detection. We selected practices that (1) span the four principles of the AI Framework, (2) reflect early adoption of AI implementation, and (3) are highly relevant for the specific use case. For each selected practice, we considered pertinent criteria from the AI Framework, National institute for Standards and Technology and OMB guidance, and relevant AI Executive Orders, along with relevant key questions for each practice. We then assessed the key questions for the Automated PII Detection use case to determine whether the practice was:[1]

- fully implemented—the agency provided evidence which showed that it fully or largely addressed the key considerations.

---

[1]Each question to consider is designed to indicate key factors associated with implementing a certain practice.

- mostly implemented— the agency provided evidence that it had
  addressed most of the key considerations.

- partially implemented—the agency provided evidence that it had
  addressed at least some of the key considerations.

- minimally implemented—the agency provided evidence that it had
  addressed at least one of the key considerations.

- not implemented—the agency did not provide evidence that it had
  addressed any of the key considerations.

For this analysis, we obtained documentation from CISA which included
technical specifications and requirements, workflows, data
characterization, and test plans for the AI, as well as documentation on
strategic and implementation plans. We also reviewed agency guidance
documents and relevant laws such as the Cybersecurity Information
Sharing Act of 2015, the Privacy Act of 1974, the Homeland Security Act
of 2002, the E-Government Act of 2002, and the Paperwork Reduction
Act. We interviewed relevant officials from CISA's Cybersecurity Division,
Office of Privacy, and Subdivision of Mission Engineering responsible for
managing Automated PII Detection on aspects related to each key
practice. For each key practice, one analyst provided justification for
ranking a key practice based on the groupings noted above and a second
analyst reviewed the justification to ensure it was sufficient.

We conducted this performance audit from September 2022 to February
2024 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 objectives.

Appendix II: Snapshot of AI in Cybersecurity

The figure below summarizes the issues participants discussed at the 2017 Comptroller General forum on Artificial Intelligence on the status of AI in cybersecurity.[1]

---

[1]GAO, *Technology Assessment: Artificial Intelligence: Emerging Opportunities, Challenges, and Implications,* GAO-18-142SP (Washington, D.C.: Mar. 28, 2018).

Figure 7: Snapshot of AI in Cybersecurity from the 2017 Comptroller General Forum on Artificial Intelligence

Snapshot

AI in Cybersecurity

This cybersecurity snapshot summarizes the issues participants discussed at the 2017 Comptroller General Forum on Artificial Intelligence.

Key Policy Areas for Consideration

- Explore options to incentivize both innovation and security in autonomous systems
- Assess the usefulness of a risk-based approach to determining if machine-learning algorithms adhere to legal requirements or ethical norms

Applications

Automated systems and advanced algorithms can help cybersecurity professionals in a variety of ways. For example, these systems can help reduce the time and effort it takes to perform key cybersecurity tasks, such as:

- identifying vulnerabilities,
- patching vulnerabilities,
- detecting attacks, and
- defending against active attacks.

Expert systems remain the most common systems used for cybersecurity, but newer approaches incorporate a combination of machine learning with human expertise to build a predictive model of cyber-attacks. As shown in the figure below, an AI system may use both unsupervised and supervised machine learning to conduct analysis of potential threats.
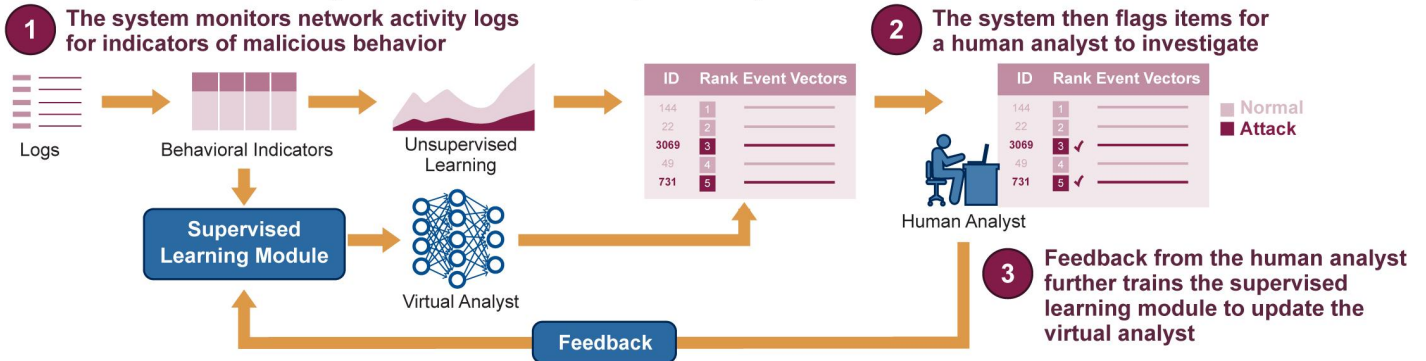
Selected Benefits

- Reducing human workload
- Increasing accuracy in detection of cyber threats
- Processing large amounts of data in short time spans

Selected Challenges

- Depending on human intervention for ongoing operation and periodic maintenance, including the identification and/or verification of attacks
- Addressing ethical and legal concerns of how AI uses personal data
- Addressing AI's own vulnerability to cyber attacks that attempt to maliciously manipulate the AI system's actions
- Countering automated or AI-based attacks

Illustration of machine learning with human feedback for cybersecurity

1. **The system monitors network activity logs for indicators of malicious behavior**

2. **The system then flags items for a human analyst to investigate**

Logs → Behavioral Indicators → Unsupervised Learning

Supervised Learning Module → Virtual Analyst

Feedback

ID	Rank	Event Vectors
144	1	
22	2	
3069	3	
49	4	
731	5	

Human Analyst

ID	Rank	Event Vectors
144	1	
22	2	
3069	3 ✓	
49	4	
731	5 ✓	

■ Normal
■ Attack

3. **Feedback from the human analyst further trains the supervised learning module to update the virtual analyst**

Source: GAO (photo); GAO, adapted from video, Veeramachaneni, Arnaldo et al., *AI²: Training a Big Data Machine to Defend* (https://www.youtube.com/watch?v=b6Hf1O_vpwQ) (illustration). | GAO-24-106246

Accessible text for Figure 7: Snapshot of AI in Cybersecurity from the 2017 Comptroller General Forum on Artificial Intelligence

<u>AI in Cybersecurity: This cybersecurity snapshot summarizes the issues participants discussed at the 2017 Comptroller General Forum on Artificial Intelligence.</u>

Key Policy Areas for Consideration

- Explore options to incentivize both innovation and security in autonomous systems
- Assess the usefulness of a risk-based approach to determining if machine- learning algorithms adhere to legal requirements or ethical norms

Applications

Automated systems and advanced algorithms can help cybersecurity professionals in a variety of ways. For example, these systems can help reduce the time and effort it takes to perform key cybersecurity tasks, such as:

- identifying vulnerabilities,
- patching vulnerabilities,
- detecting attacks, and
- defending against active attacks.

Expert systems remain the most common systems used for cybersecurity, but newer approaches incorporate a combination of machine learning with human expertise to build a predictive model of cyber-attacks. As shown in the figure below, an AI system may use both unsupervised and supervised machine learning to conduct analysis of potential threats.

Selected Benefits

- Reducing human workload
- Increasing accuracy in detection of cyber threats
- Processing large amounts of data in short time spans

GAO-24-106246 Artificial Intelligence

Selected Challenges

- Depending on human intervention for ongoing operation and periodic maintenance, including the identification and/or verification of attacks

- Addressing ethical and legal concerns of how AI uses personal data

- Addressing AI's own vulnerability to cyber attacks that attempt to maliciously manipulate the AI system's actions

- Countering automated or AI-based attacks

Illustration of machine learning with human feedback for cybersecurity

1) The system monitors network activity logs for indicators of malicious behavior

    a) Logs

    b) Behavioral indicators

        i) Unsupervised learning

        ii) Supervised learning modules

        iii) Virtual analyst

        iv) Feedback

2) The system then flags items for a human analyst to investigate

3) Feedback from the human analyst further trains the supervised learning module to update the virtual analyst

Source: GAO (photo); GAO, adapted from video, Veeramachaneni, Arnaldo et al., AI2: Training a Big Data Machine to Defend (https://www.youtube.com/watch?v=b6Hf1O_vpwQ) (illustration). | GAO-24-106246

Appendix III: Comments from the Department of Homeland Security

**Homeland
Security**

January 17, 2024

Candice N. Wright
Director, Science, Technology Assessment, and Analytics
U.S. Government Accountability Office
441 G Street, NW
Washington, DC 20548-0001

Kevin Walsh
Director, Information Technology and Cybersecurity
U.S. Government Accountability Office
441 G Street, NW
Washington, DC 20548-0001

Re: Management Response to Draft Report GAO-24-106246 "ARTIFICIAL
     INTELLIGENCE: Fully Implementing Key Practices Could Help DHS Ensure
     Responsible AI Use for Cybersecurity"

Dear Ms. Wright and Mr. Walsh,

Thank you for the opportunity to comment on this draft report. The U.S. Department of
Homeland Security (DHS or the Department) appreciates the U.S. Government
Accountability Office's (GAO) work in planning and conducting its review and issuing
this report.

DHS leadership is pleased to note GAO's recognition that DHS established an inventory
of Artificial Intelligence (AI) use cases to promote transparency and inform the public
about how AI is being used. GAO also acknowledged DHS's implementation of key AI
Framework practices in managing and overseeing its use of AI for cybersecurity. DHS
remains committed to ensuring responsible use of AI within the Department.

As part of this commitment, the Cybersecurity and Infrastructure Security Agency
(CISA) will continue work with the DHS Office of the Chief Information Officer Chief
Technology Officer Directorate (CTOD) to expand and improve the review process for
its AI inventory and implementation of the AI framework practices ensuring future
systems appropriately follow the governance, data, performance, and monitoring
guidance for the responsible use of AI systems. As the Automated Personally

Identifiable Information (PII) Detection component was deployed as part of the
Automated Indicator Sharing (AIS) capability before the AI Framework was issued, the
CISA Cybersecurity Division (CSD) Mission Engineering (ME) Subdivision will
continue to work DHS CTOD and the CISA Chief Data Officer to determine the best
approach to further apply key AI Framework practices in accordance with Office of
Management and Budget and DHS guidance to AI components deployed in production
systems to ensure accountable and responsible use of AI.

The draft report contained eight recommendations with which the Department concurs.
Enclosed find our detailed response to each recommendation. DHS will submit technical
comments addressing several accuracy, contextual, and other issues under a separate
cover for GAO's consideration.

Again, thank you for the opportunity to review and comment on this draft report. Please
feel free to contact me if you have any questions. We look forward to working with you
again in the future.

Sincerely,

JIM H
CRUMPACKER

Digitally signed by JIM H
CRUMPACKER
Date: 2024.01.17 16:12:15 -05'00'

JIM H. CRUMPACKER, CIA, CFE
Director
Departmental GAO-OIG Liaison Office

Enclosure

2

**Enclosure: Management Response to Recommendations
Contained in GAO-24-106246**

GAO recommended that the Chief Technology Officer:

Recommendation 1: Expand its review process to include steps to verify the accuracy
of its AI inventory submissions.

Response: Concur. The CTOD agrees with the importance of improving the process of
compiling and maintaining the DHS AI inventory, and notes that major improvements
have already been implemented. For example, CTOD launched a centralized internal
repository in June 2023 which provides interim updates to the public inventory. CTOD
also updated the process for reviewing and adding AI Use Cases in Fall 2023, and
implemented new oversight mechanisms to identify new AI Use Cases, such as
expanding the central internal repository to include new additions that then undergo the
new process for review and approval. CTOD will further update the review process and
evaluation criteria to ensure it is standardized and rigorous. Estimated Completion Date
(ECD): May 31, 2024.

GAO recommended that the Director of CISA:

Recommendation 2: Develop metrics to consistently measure progress towards all
stated goals and objectives for Automated PII Detection.

Response: Concur. CISA's CSD ME Subdivision, in collaboration with CISA's Privacy
Officer and other CISA stakeholders, will determine a method for documenting and
tracking metrics to measure the goals and objectives for Automated PII Detection and
will develop metrics to consistently measure the stated goals and objectives for
Automated PII Detection. ECD: September 30, 2024.

Recommendation 3: Clearly define the roles and responsibilities and delegation of
authority of all relevant stakeholders involved in managing and overseeing the
implementation of the Automated PII Detection component to ensure effective operations
and sustained oversight.

Response: Concur. In coordination with CSD leadership and the CSD Technology
Director, the CISA Privacy Officer, and other CISA stakeholders, CISA CSD ME will
develop a document that defines the roles and responsibilities of the stakeholders
involved in managing and overseeing the implementation of the Automated PII Detection
component. This work product will augment existing documentation that articulates
stakeholder roles and responsibilities for CISA CSD ME's engineering process. ECD:
September 30, 2024.

3

Recommendation 4: Document the sources and origins of data used to develop the Automated PII Detection component.

Response: Concur. CISA CSD ME will document the sources and origins of data used to develop the Automated PII Detection component. ECD: September 30, 2024.

Recommendation 5: Take steps to assess and document the reliability of data used to enhance the representativeness, quality, and accuracy of the Automated PII Detection component.

Response: Concur. CISA CSD ME will assess and document the reliability of data used to enhance the representativeness, quality, and accuracy of the Automated PII Detection component. ECD: September 30, 2024.

Recommendation 6: Document its process for optimizing the elements used within the Automated PII Detection component.

Response: Concur. CISA CSD ME will develop a document that defines the process for optimizing the elements used with the Automated PII Detection component. ECD: September 30, 2024.

Recommendation 7: Document its methods for testing performance including limitations, and corrective actions taken to minimize undesired effects of the Automated PII Detection component to ensure transparency about the system's performance.

Response: Concur. CISA CSD ME currently conducts various forms of testing throughout the engineering lifecycle to include manual and automated testing. CISA CSD ME will develop test planning documentation to further define the methods for testing performance of the Automate PII Detection component. Results of testing will continue to be issued via test reports, which document limitations and corrective actions taken to minimize the undesired effects of the Automate PII Detection component of AIS. ECD: December 31, 2024.

Recommendation 8: Establish specific procedures and frequencies to monitor Automated PII Detection component to ensure it performs as intended.

Response: Concur. CISA CSD ME, in coordination with the CSD Technology Director, the CISA Privacy Officer, and other CISA stakeholders will update its AIS monitoring plans to include procedures and frequencies specific to monitoring the Automated PII Detection component. ECD: December 31, 2024.

4

Accessible text for Appendix III: Comments from the Department of Homeland Security

January 17, 2024

U.S. Department of Homeland Security
Washington, DC 20528

Candice N. Wright
Director, Science, Technology Assessment, and Analytics
U.S. Government Accountability Office 441 G Street, NW
Washington, DC 20548-0001

Kevin Walsh

Director, Information Technology and Cybersecurity

U.S. Government Accountability Office 441 G Street, NW

Washington, DC 20548-0001

Re: Management Response to Draft Report GAO-24-106246 "ARTIFICIAL INTELLIGENCE: Fully Implementing Key Practices Could Help DHS Ensure Responsible AI Use for Cybersecurity"

Dear Ms. Wright and Mr. Walsh,

Thank you for the opportunity to comment on this draft report. The U.S. Department of Homeland Security (DHS or the Department) appreciates the U.S. Government Accountability Office's (GAO) work in planning and conducting its review and issuing this report.

DHS leadership is pleased to note GAO's recognition that DHS established an inventory of Artificial Intelligence (AI) use cases to promote transparency and inform the public about how AI is being used. GAO also acknowledged DHS's implementation of key AI Framework practices in managing and overseeing its use of AI for cybersecurity. DHS remains committed to ensuring responsible use of AI within the Department.

As part of this commitment, the Cybersecurity and Infrastructure Security Agency (CISA) will continue work with the DHS Office of the Chief Information Officer Chief

Technology Officer Directorate (CTOD) to expand and improve the review process
for its AI inventory and implementation of the AI framework practices ensuring future
systems appropriately follow the governance, data, performance, and monitoring
guidance for the responsible use of AI systems. As the Automated Personally
Identifiable Information (PII) Detection component was deployed as part of the
Automated Indicator Sharing (AIS) capability before the AI Framework was issued,
the CISA Cybersecurity Division (CSD) Mission Engineering (ME) Subdivision will
continue to work DHS CTOD and the CISA Chief Data Officer to determine the best
approach to further apply key AI Framework practices in accordance with Office of
Management and Budget and DHS guidance to AI components deployed in
production systems to ensure accountable and responsible use of AI.

The draft report contained eight recommendations with which the Department
concurs. Enclosed find our detailed response to each recommendation. DHS will
submit technical comments addressing several accuracy, contextual, and other
issues under a separate cover for GAO's consideration.

Again, thank you for the opportunity to review and comment on this draft report.
Please feel free to contact me if you have any questions. We look forward to working
with you again in the future.

Sincerely,

JIM H. CRUMPACKER, CIA, CFE

Director

Departmental GAO-OIG Liaison Office

Enclosure: Management Response to Recommendations Contained in GAO-24-106246

GAO recommended that the Chief Technology Officer:

**Recommendation 1: Expand its review process to include steps to verify the
accuracy of its AI inventory submissions.**

**Response: Concur. The CTOD agrees with the importance of improving the
process of compiling and maintaining the DHS AI inventory, and notes that
major improvements have already been implemented. For example, CTOD
launched a centralized internal repository in June 2023 which provides interim
updates to the public inventory. CTOD also updated the process for reviewing**

and adding AI Use Cases in Fall 2023, and implemented new oversight
mechanisms to identify new AI Use Cases, such as expanding the central
internal repository to include new additions that then undergo the new process
for review and approval. CTOD will further update the review process and
evaluation criteria to ensure it is standardized and rigorous. Estimated
Completion Date (ECD): May 31, 2024.

GAO recommended that the Director of CISA:

**Recommendation 2: Develop metrics to consistently measure progress
towards all stated goals and objectives for Automated PII Detection.**

**Response: Concur. CISA's CSD ME Subdivision, in collaboration with CISA's
Privacy Officer and other CISA stakeholders, will determine a method for
documenting and tracking metrics to measure the goals and objectives for
Automated PII Detection and will develop metrics to consistently measure the
stated goals and objectives for Automated PII Detection. ECD: September 30,
2024.**

**Recommendation 3: Clearly define the roles and responsibilities and
delegation of authority of all relevant stakeholders involved in managing and
overseeing the implementation of the Automated PII Detection component to
ensure effective operations and sustained oversight.**

**Response: Concur. In coordination with CSD leadership and the CSD
Technology Director, the CISA Privacy Officer, and other CISA stakeholders,
CISA CSD ME will develop a document that defines the roles and
responsibilities of the stakeholders involved in managing and overseeing the
implementation of the Automated PII Detection component. This work product
will augment existing documentation that articulates stakeholder roles and
responsibilities for CISA CSD ME's engineering process. ECD: September 30,
2024.**

**Recommendation 4: Document the sources and origins of data used to
develop the Automated PII Detection component.**

**Response: Concur. CISA CSD ME will document the sources and origins of
data used to develop the Automated PII Detection component. ECD:
September 30, 2024.**

**Recommendation 5: Take steps to assess and document the reliability of data
used to enhance the representativeness, quality, and accuracy of the
Automated PII Detection component.**

**Response: Concur. CISA CSD ME will assess and document the reliability of
data used to enhance the representativeness, quality, and accuracy of the
Automated PII Detection component. ECD: September 30, 2024.**

**Recommendation 6: Document its process for optimizing the elements used
within the Automated PII Detection component.**

**Response: Concur. CISA CSD ME will develop a document that defines the
process for optimizing the elements used with the Automated PII Detection
component. ECD: September 30, 2024.**

**Recommendation 7: Document its methods for testing performance including
limitations, and corrective actions taken to minimize undesired effects of the
Automated PII Detection component to ensure transparency about the
system's performance.**

**Response: Concur. CISA CSD ME currently conducts various forms of testing
throughout the engineering lifecycle to include manual and automated testing.
CISA CSD ME will develop test planning documentation to further define the
methods for testing performance of the Automate PII Detection component.
Results of testing will continue to be issued via test reports, which document
limitations and corrective actions taken to minimize the undesired effects of
the Automate PII Detection component of AIS. ECD: December 31, 2024.**

**Recommendation 8: Establish specific procedures and frequencies to monitor
Automated PII Detection component to ensure it performs as intended.**

**Response: Concur. CISA CSD ME, in coordination with the CSD Technology
Director, the CISA Privacy Officer, and other CISA stakeholders will update its
AIS monitoring plans to include procedures and frequencies specific to
monitoring the Automated PII Detection component. ECD: December 31, 2024.**

Appendix IV: GAO Contacts and Staff Acknowledgments

GAO Contacts

Candice N. Wright at (202) 512-6888 or WrightC@gao.gov
Kevin Walsh at (202) 512-6151 or WalshK@gao.gov

Staff Acknowledgments

In addition to the contacts named above, principal contributors to this report were Farahnaaz Khakoo-Mausel (Assistant Director), Michael Walton (Analyst-in-Charge from August 26, 2023), Jon Menaster (Analyst-in-Charge until August 26, 2023), Taka Ariga, Nicole Catanzarite, Jehan Chase, Jillian Clouse, Bill Cook, Louise Fickel, Alexander Gromadzki, Paris Hall, Ryan Han, Andrew Kurtzman, Stephanie Palmer, Andrew Stavisky, and Jessica Steele.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through our website. Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. You can also subscribe to GAO's email updates to receive notification of newly posted products.

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, https://www.gao.gov/ordering.htm.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on Facebook, Flickr, Twitter, and YouTube.
Subscribe to our RSS Feeds or Email Updates. Listen to our Podcasts.
Visit GAO on the web at https://www.gao.gov.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact FraudNet:

Website: https://www.gao.gov/about/what-gao-does/fraudnet

Automated answering system: (800) 424-5454 or (202) 512-7700

Congressional Relations

A. Nicole Clowers, Managing Director, ClowersA@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

Stephen J. Sanford, Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814,
Washington, DC 20548