TREASURY INSPECTOR GENERAL FOR TAX ADMINISTRATION

# *The First Phase of the Data Loss Prevention Solution Is Working As Intended, but the Remaining Phases Continue to Experience Delays*

**August 22, 2019**

**Reference Number: 2019-20-049**

This report has cleared the Treasury Inspector General for Tax Administration disclosure review process and information determined to be restricted from public release has been redacted from this document.

*Phone Number / 202-622-6500*
*E-mail Address / TIGTACommunications@tigta.treas.gov*
*Website / http://www.treasury.gov/tigta*


**To report fraud, waste, or abuse, call our toll-free hotline at:**

1-800-366-4484

**By Web:**

***www.treasury.gov/tigta/***

**Or Write:**

Treasury Inspector General for Tax Administration
P.O. Box 589
Ben Franklin Station
Washington, D.C. 20044-0589

Information you provide is confidential and you may remain anonymous.

**THE FIRST PHASE OF THE DATA LOSS PREVENTION SOLUTION IS WORKING AS INTENDED, BUT THE REMAINING PHASES CONTINUE TO EXPERIENCE DELAYS**

# Highlights

**Final Report issued on August 22, 2019**

Highlights of Reference Number: 2019-20-049 to the Commissioner of Internal Revenue.

## IMPACT ON TAXPAYERS

The IRS is entrusted with protecting information received from taxpayers, including Personally Identifiable Information and tax account data. Allowing this information to be removed or exfiltrated for unauthorized purposes could erode public trust in the IRS's ability to administer our Nation's tax system and in the voluntary compliance nature of tax filing.

## WHY TIGTA DID THE AUDIT

This audit was initiated to determine whether the IRS properly implemented controls to prevent data loss, including data exfiltration of personal information. The IRS is implementing a Data Loss Prevention software solution to identify and prevent Personally Identifiable Information from leaving the IRS network, whether intentionally or unintentionally. The software has multiple components that are being implemented over several years, and this audit evaluated the progress of the implementation.

## WHAT TIGTA FOUND

The Safeguarding Personally Identifiable Information Data Extracts Project, which is responsible for implementing the Data Loss Prevention solution, started in Calendar Year 2010 and is ongoing. The project team implemented and expanded the Data-in-Motion component of the solution that includes reviewing unencrypted e-mail and attachments, file transfers, and web traffic for the most common types of Personally Identifiable Information used by the IRS. Our testing indicated that the Data-in-Motion component generally identified and blocked common Personally Identifiable Information types from exfiltration by e-mail as designed, and that potential incidents identified by the solution were reviewed and resolved correctly. However, continued delays with implementing other components are preventing realization of the full benefits of the Data Loss Prevention solution.

The causes of the delays include technical, project management, and administrative issues. Because of the delays, two key components involving data in repositories and data in use are still not operational more than eight years after the project started. Without these components, Personally Identifiable Information continues to be at risk of loss. The delays have also resulted in the inefficient use of resources of approximately $1.2 million in software costs for the components that are not operational.

## WHAT TIGTA RECOMMENDED

TIGTA recommended that the Chief Information Officer deploy the components of the Data Loss Prevention solution, ensure that project documents are prepared and maintained as required, and ensure that any issues requiring negotiations with the National Treasury Employees Union are identified and negotiations started promptly.

The IRS agreed with all three recommendations and plans to deploy the remaining components of the Data Loss Prevention solution and ensure that project documents are consistently prepared and maintained during the deployment of the remaining components. In addition, the IRS stated that the Memorandum of Understanding with the National Treasury Employees Union is currently in the process of concurrence signatures, and the IRS plans to notify the Union of any issues regarding the production implementation of the remaining components.

TREASURY INSPECTOR GENERAL
FOR TAX ADMINISTRATION

DEPARTMENT OF THE TREASURY
WASHINGTON, D.C. 20220

August 22, 2019

**MEMORANDUM FOR** COMMISSIONER OF INTERNAL REVENUE

**FROM:** Michael E. McKenney
Deputy Inspector General for Audit

**SUBJECT:** Final Audit Report – The First Phase of the Data Loss Prevention Solution Is Working As Intended, but the Remaining Phases Continue to Experience Delays (Audit # 201820003)

This report presents the results of our review to evaluate whether the Internal Revenue Service (IRS) has properly implemented controls to prevent data loss, including data exfiltration of personal information. This audit is included in our Fiscal Year 2019 Annual Audit Plan and addresses the major management challenge of Security Over Taxpayer Data and Protection of IRS Resources.

Management's complete response to the draft report is included as Appendix V.

Copies of this report are also being sent to the IRS managers affected by the report recommendations. If you have any questions, please contact me or Danny R. Verneuille, Assistant Inspector General for Audit (Security and Information Technology Services).

# *Table of Contents*

Background ..... Page 1

Results of Review ..... Page 6

The Data-in-Motion Component of the Data Loss Prevention Solution Has Been Implemented and Is Working As Intended ..... Page 6

Delays Are Preventing the IRS From Realizing All the Benefits of the Data Loss Prevention Solution ..... Page 9

Recommendations 1 through 3: ..... Page 16

**Appendices**

Appendix I – Detailed Objective, Scope, and Methodology ..... Page 17

Appendix II – Major Contributors to This Report ..... Page 19

Appendix III – Report Distribution List ..... Page 20

Appendix IV – Outcome Measure ..... Page 21

Appendix V – Management's Response to the Draft Report ..... Page 22

Appendix VI – Office of Audit Comments on Management's Response ..... Page 26

# Abbreviations

| | |
|---|---|
| DAR | Data-at-Rest |
| DIM | Data-in-Motion |
| DIU | Data-in-Use |
| DLP | Data Loss Prevention |
| IRS | Internal Revenue Service |
| NTEU | National Treasury Employees Union |
| OMB | Office of Management and Budget |
| PII | Personally Identifiable Information |
| SPIIDE | Safeguarding Personally Identifiable Information Data Extracts |
| SSN | Social Security Number |
| TIGTA | Treasury Inspector General for Tax Administration |

# Background

All Federal Government agencies have the fiduciary responsibility to safeguard information in their possession and prevent its loss to earn and retain the trust of the American public. The importance of protecting these data is reflected in the various statutes and departmental and agency guidance specific to data protection and privacy.

- The Privacy Act of 1974[1] requires agencies to establish appropriate administrative, technical, and physical safeguards to ensure the security and confidentiality of records and to protect against any anticipated threats or hazards to the records' security or integrity that could result in substantial harm, embarrassment, inconvenience, or unfairness to any individual on whom information is maintained.
- The Office of Management and Budget (OMB) has released several memoranda to Federal agencies to address protecting the vast quantities of Personally Identifiable Information (PII) managed by the Federal Government. These include OMB M-06-16, *Protection of Sensitive Agency Information*, dated June 2016, which provides guidance on protecting data extracts containing PII, and OMB M-17-12, *Preparing for and Responding to a Breach of Personally Identifiable Information*, dated January 2017, which sets forth policy for Federal agencies to prepare for and respond to a breach of PII.
- The Department of the Treasury's (hereafter referred to as the Treasury Department) Office of the Chief Information Officer instituted additional controls in its memorandum M-09-04, related to the management of the Treasury Department's cybersecurity environment. These new controls focused particularly on the storage of data on removable media and the unauthorized transmission of information outside the Treasury Department, *i.e.*, data exfiltration. Implementation of these controls was intended to help ensure better protection from emerging threats.
- National Institute of Standards and Technology Special Publication 800-122, *Guide to Protecting the Confidentiality of Personally Identifiable Information*,[2] and Special Publication 800-53 revision 4, *Security and Privacy Controls for Federal Information Systems and Organizations*,[3] recommend that agencies implement automated tools, such as a network data leakage prevention tool, to monitor transfers of PII and to monitor inbound and outbound communications for unauthorized activities.

[1] 5 U.S.C. § 552a (2013).
[2] National Institute of Standards and Technology Special Publication 800-122, *Guide to Protecting the Confidentiality of Personally Identifiable Information* (Apr. 2010).
[3] National Institute of Standards and Technology Special Publication 800-53 Revision 4, *Security and Privacy Controls for Federal Information Systems and Organizations* (Apr. 2013).

The Internal Revenue Service (IRS) is entrusted with protecting information received from taxpayers, including PII and tax account data. PII is any information that, by its nature or in combination with other information, may be used to uniquely identify an individual. Examples include name, Social Security Number (SSN), date and place of birth, mother's maiden name, and biometric records. Allowing PII to be removed or exfiltrated for unauthorized purposes could erode public trust in the IRS's ability to administer our Nation's tax system and in the voluntary compliance nature of tax filing. As part of this responsibility, the IRS relies on its employees to safeguard tax information and implement systemic controls and measures to safeguard such information.

In an effort to better control and protect information, such as PII, technical solutions known as Data Loss Prevention (DLP), also referred to as Data Leak Prevention, are available to organizations. DLP is the practice of detecting and preventing confidential data, such as PII, from being "leaked" out of an organization's boundaries, either intentionally or unintentionally.

In a white paper on DLP,[4] Ernst and Young wrote:

> *In addition to obvious data loss methods such as the loss of physical assets such as laptops, many data loss incidents are due to accidental disclosure through electronic transmissions. In most cases, end users do not realize the risks associated with sending sensitive data through unencrypted e-mail, instant messages, webmail, and file transfer tools.*

While DLP software solutions can vary in their capabilities, they commonly have the ability to intercept some malicious or criminal attempts to steal information. An important distinction between DLP and other security technologies is that it focuses on identifying sensitive information that is critical to an organization and may be at risk by personnel who are authorized to access the information (and others). This is in contrast with more traditional efforts such as using a firewall or an intrusion detection system to prevent unauthorized access to data. Unauthorized activities by employees or contractors to cause harm (wittingly or unwittingly) are known by the term 'insider threat.' Insiders are considered one of today's biggest security threats across the government and commercial sectors, and therefore, some kind of DLP capability is essential to reduce risks. However, given the complexities of identifying and preventing these activities, it is understood that DLP technology alone cannot identify and prevent all methods of data theft.

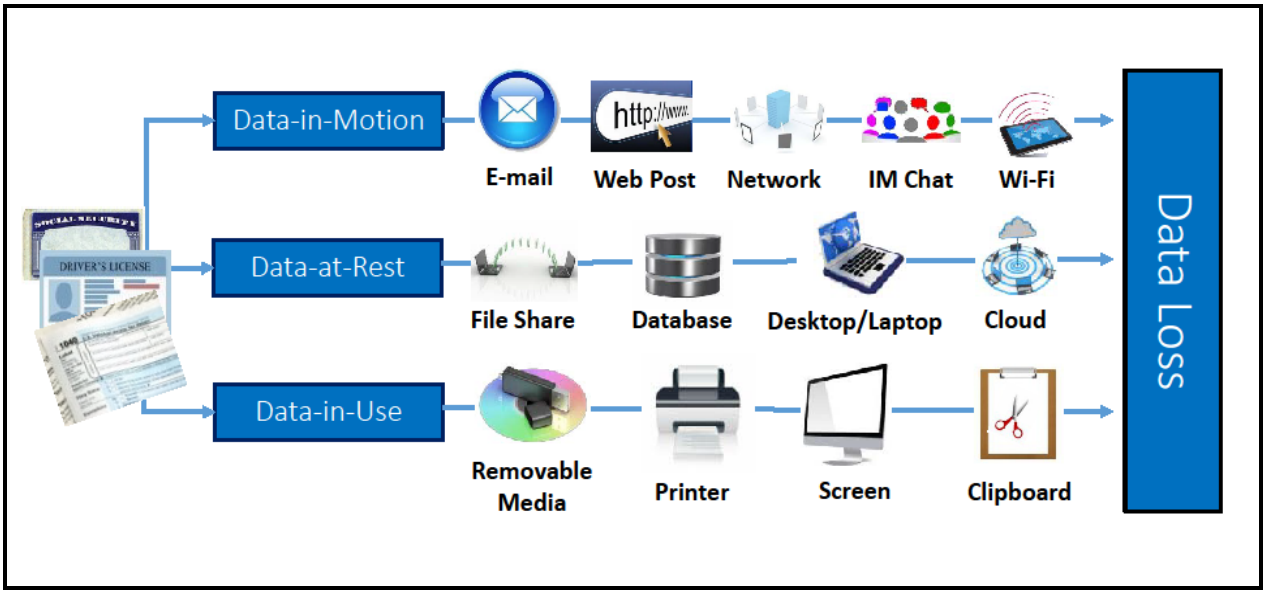
An example of how easily sensitive data can be compromised and misused by insiders involves two well-known companies developing self-driving car technology. An employee of Google's self-driving car division (now Waymo) downloaded thousands of files with proprietary information by a simple file transfer to a USB drive. The employee then quit and started his own company specializing in self-driving cars, which was shortly after acquired by Uber, a direct

[4] Ernst and Young, *Data Loss Prevention: Keeping Your Sensitive Data Out of the Public Domain* (Oct. 2011).

competitor of Waymo. It was then alleged that Uber used the information for its benefit.[5] If DLP software had been used and properly configured, this incident may have been prevented or readily identified to minimize the damage.

DLP capability is generally broken down into the protection of three key types of data as listed in Figure 1. To be considered a full solution, a DLP solution must have the capability to address all three data types and be integrated by a centralized management function.

***Figure 1: The Potential Data Loss Sources by Type of Data***




*Source: Treasury Inspector General for Tax Administration (TIGTA) figure based on DLP definitions.*

The Data-in-Motion (DIM) component of DLP covers data being transmitted outside of the organization through Internet routers, e-mail gateways, and web proxies. This includes data being transmitted through e-mail, Internet chat, and information entered into web pages. The Data-at-Rest (DAR) component covers data residing in enterprise data repositories. This includes data files, file servers, storage area networks, and even end-user workstation hard disks. The Data-in-Use (DIU) component covers data accessed or used by a system at a point in time. This includes copying data to a thumb drive, sending information to a printer, or even cutting and pasting between applications.

In response to OMB and Treasury Department guidance, the IRS created the Safeguarding Personally Identifiable Information Data Extracts (SPIIDE) Project to oversee the implementation of controls over data loss, specifically the DLP solution, in Calendar Year 2010.

[5] YHB CPAs and Consultants, *A Case for Data Loss Prevention Tools* (Mar. 2017).

The DLP solution, as envisioned for use by the IRS, was to be deployed in several releases to accomplish the following tasks.

1. Monitor DIM across the IRS infrastructure perimeter, and based on IRS policy, allow or prevent PII from leaving the IRS infrastructure or make the user confirm the transmission of the data.
2. Discover DAR residing across the IRS infrastructure and assess whether PII has adequate protections.
3. Monitor DIU created and manipulated on users workstations and allow or prevent the distribution of the data.

The IRS Cybersecurity Architecture and Implementation group is responsible for the SPIIDE Project, with the project team being responsible for the technical development, deployment, implementation, and testing of the DLP solution based on commercial off-the-shelf software. The SPIIDE Project is governed by an Executive Steering Committee with oversight by the Management Level Governance Board. There is also a dedicated working group that was established to monitor the DLP program's effectiveness, provide input on emergent decision points, and ensure that the right resources are involved to drive DLP success. The working group includes members from various IRS functional areas.

To accomplish the stated goals of the SPIIDE Project, the IRS contracted with a third-party vendor for DLP software licenses in Calendar Year 2011. Initially, the IRS planned to fully implement the DLP solution in April 2012; however, this date was later changed to July 2012. In April 2013, the SPIIDE Project Executive Steering Committee approved changing the implementation date of the DLP solution to December 31, 2014.

In our September 2014 audit report[6] on the progress of the SPIIDE Project's implementation of the DLP solution, TIGTA reported that the SPIIDE Project team had completed key required enterprise life cycle deliverables and had identified and addressed security weaknesses as they were detected. However, the report also indicated that the SPIIDE Project team continued to face challenges with timely implementing the DLP solution to protect disclosure of PII and other data.

This review was performed with information obtained from the Information Technology organization's Cybersecurity office in New Carrollton, Maryland, during the period October 2017 through February 2019. We conducted this performance audit in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objective. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objective.

[6] TIGTA, Ref. No. 2014-20-087, *While the Data Loss Prevention Solution Is Being Developed, Stronger Oversight and Process Enhancements Are Needed for Timely Implementation Within Budget* (Sept. 2014).

Detailed information on our audit objective, scope, and methodology is presented in Appendix I. Major contributors to the report are listed in Appendix II.

# Results of Review

## The Data-in-Motion Component of the Data Loss Prevention Solution Has Been Implemented and Is Working As Intended

The DIM component of the IRS DLP solution was deployed to a production environment in May 2015. The current screening process includes reviewing unencrypted e-mail and attachments, file transfers, and web traffic for the most common types of PII used by the IRS. Our testing indicated that the DIM component generally identified and blocked common PII types from exfiltration by e-mail as designed and that potential incidents identified by the DLP solution were reviewed and resolved as required.

### The DIM component criteria is based on PII

For internal e-mails with PII, the IRS uses the Secure Enterprise Messaging System, which enables the IRS to digitally encrypt e-mail messages and attachments sent between IRS employees. Accordingly, IRS policy states that employees may not e-mail PII outside the IRS unless there is an approved exception from the Information Technology organization.[7] The DIM component is a control to ensure that this policy is followed, and it is currently the only operational component of the DLP solution. If the DIM component identifies potential unencrypted PII leaving the internal network, it will take action to prevent exfiltration. For example, an e-mail meeting specific criteria found in the DLP system policies will be blocked, and the sending employee will receive a warning e-mail explaining why his or her e-mail was not sent. The e-mail will also be manually reviewed to determine if the circumstances were suspicious and if further action is warranted.

The IRS initially focused its DLP system policies on SSNs. When the DIM component was implemented, the first set of policies included rules to identify SSNs based on a specific pattern and in association with certain keywords, such as "Social Security Number" or "SSN." After addressing SSNs, the next set of policies implemented was focused on identifying password-related terms. The password policy is comprised of multiple detection rules, including detection of employee identifiers in combination with common password-related keywords. It also includes detection of common IRS application names in proximity to password-related keywords.

[7] The IRS is testing a type of secure communication with taxpayers through the Taxpayer Digital Communications program. This provides a way for selected taxpayers and their representatives to exchange secure messages with IRS employees for a variety of reasons, including providing requested documentation for examinations.

Subsequently, the IRS expanded the SSN policy set to include similar types of PII that the IRS assigns internally for other tax purposes, such as Taxpayer Identification Numbers and Adoption Taxpayer Identification Numbers. The format of SSNs and these internal numbers are similar, consisting of nine digit number strings, so there are similar complications with identifying them. For example, when the IRS first began developing the policies used to identify SSNs, it discovered other types of common nine digit numbers, such as a 5+4 zip code, that could also be considered a match. Accordingly, the IRS has been working towards refining the formats and types of data to be included in the DIM component to increase coverage while also decreasing the number of false positives. This includes creating exclusion rules that will allow data proven to result in a false positive to pass through the system. Over time, the IRS has further expanded the types of data for review, with the most recent policy implemented to identify and protect unencrypted credit card numbers. Overall, the DIM component is designed to prevent employees from sending unencrypted e-mails with various PII types to external parties, whether intentionally or unintentionally.

## The DIM component generally identified and blocked common types of PII

To test whether the DIM component was working as intended, we created test e-mails containing various types of common PII that were supposed to be identified and blocked based on current DLP policies. We created 30 test e-mails that contained examples of test PII related to 12 common policy rules, with multiple variations of some rules, *e.g.*, PII in the body of the e-mail versus PII included in an e-mail attachment.

We coordinated with an IRS employee, who attempted to send the test e-mails to an external e-mail address, and we documented in real time whether the DLP solution blocked the unencrypted e-mails containing PII based on the related PII policy rule. We then researched any e-mails that were not blocked to determine whether the criteria should be updated or whether there was an acceptable reason the e-mails were not blocked. Based on this determination and for the specific policy rules tested (including certain Taxpayer Identification Numbers, keywords, and password terms), the DLP solution identified and blocked our test e-mails as expected.

## Potential incidents identified by the DLP solution were reviewed and resolved properly

When the DLP solution identifies and blocks data that meet certain exception criteria as designated in DLP system policies, the identified traffic is referred to the DLP Operations team, which receives the event information in the form of potential incidents. The DLP Operations team analyzes each potential incident to determine if a PII disclosure or attempted disclosure occurred. If a potential incident is confirmed, the DLP Operations team escalates it to one or more parties, depending on its categorization.

- Business System Process Liaison Event Responders are the primary recipients of DLP event alerts. They receive alerts of blocked events when an employee from the business

unit attempts to send unencrypted PII to another governmental recipient or taxpayer representative while performing his or her work. The Business System Process Liaison Event Responders are supposed to triage the events and, if necessary, contact the manager of the employee for additional information or to provide feedback.

- TIGTA receives notice of suspicious e-mails in which an employee attempts to send PII to a private or foreign e-mail address with no context in the e-mail to indicate or establish if the e-mail was work related.
- The Privacy, Governmental Liaison, and Disclosure office receives notice of all events that involve possible loss of sensitive data from a privacy perspective, such as unblocked e-mails containing SSN information. If necessary, it will contact the appropriate Business System Partner for assistance with determining if an unauthorized disclosure occurred.
- The Computer Security Incident Response Center receives notice of events that could substantially increase the risk of exposure to IRS systems. For example, employees attempting to send documents that contain network diagrams or Internet Protocol addresses[8] of information technology assets.

If the party receiving the potential incident disagrees with the initial assessment, the potential incident should be sent back to the DLP Operations team for further review.

To test whether potential incidents were reviewed, classified, and referred or otherwise resolved correctly, we reviewed a sample of incidents on the DLP system. We selected a judgmental sample[9] of 56 incidents of the 1,561 incidents that were generated during the two-week period from May 1 through May 14, 2018. Incidents are classified by a Severity Rating of High, Medium, or Low, determined by the policy set in use. For example, the severity for the SSN policy is based on the number of criteria matches identified for that incident. For each of the 14 days, we judgmentally selected two High-, one Medium-, and one Low-Severity incident (*i.e.*, four per day)[10] and reviewed them to determine whether they were triaged and remediated in accordance with IRS policy. We determined that all of the sampled incidents were reviewed and resolved correctly.

[8] An identifier for a computer or device on a suite of communication protocols used to connect hosts on the Internet. The format of an Internet Protocol address is a 32-bit numeric address written as four numbers separated by periods. Each number can be zero to 255.
[9] A judgmental sample is a nonprobability sample, the results of which cannot be used to project to the population.
[10] For two of the days, there were no Medium-Severity incidents generated. In their place, we substituted two High-Severity incidents selected at random that were generated the same days. Therefore, the total incidents by category were 30 High-, 12 Medium-, and 14 Low-Severity incidents, for a total of 56 incidents reviewed.

While we found that the DIM component of the DLP was working as intended, we are concerned with the IRS's progress on the remaining two components of the DLP solution, the DAR and DIU.

## Delays Are Preventing the IRS From Realizing All the Benefits of the Data Loss Prevention Solution

As previously stated, the IRS initiated the SPIIDE Project to help meet the various requirements set forth by the OMB and the Treasury Department and to address known shortcomings related to PII protection. In the SPIIDE Project Charter Version 1.1, dated August 31, 2011, the IRS acknowledged that it had no comprehensive plan in place to:

- Accurately discover or prevent Sensitive But Unclassified/PII data leakage.
- Ensure the public's trust of conducting business electronically, *e.g.*, electronic filing.
- Prevent unauthorized disclosure of information.

Accordingly, the intent of the SPIIDE Project was to implement a DLP solution to reduce the risk of disclosure of PII or other tax information by monitoring PII data in the IRS network. Some specific benefits of the DLP solution cited by the IRS are as follows:

- Protecting taxpayer data by reinforcing security due diligence with continuous monitoring and prevention of unauthorized or accidental use or disclosure of sensitive data.
- Protecting IRS employees by enhancing user security awareness with real-time educational notification prompts reinforcing IRS processes and policies.
- Providing enhanced protection against insider threats.
- Aiding potential investigations (including those by TIGTA) with logged event details.

The DLP solution was to be deployed in a multiple release approach utilizing a commercial off-the-shelf DLP software with the three components, DIM, DAR, and DIU, to be released in succession as shown in Figure 2.

**Figure 2: IRS Proposed DLP Release Plan**

| Release Number | Component in Release | Target Release Date |
|---|---|---|
| **1** | DIM | December 31, 2012 |
| **2** | DAR | August 1, 2013 |
| **3** | DIU | December 31, 2014 |

*Source: IRS SPIIDE Project Charter Version 1.1, dated August 31, 2011.*

However, the IRS was unable to meet any of these targeted release dates. In addition, the DAR and DIU components have still not been deployed at this time, which is more than five years after the initial target release date for the DAR component. Continued delays have prevented the risks of PII being inadvertently or intentionally released during the course of normal duties from being fully addressed and the full benefits of the DLP solution from being realized. Until all three components are operational, the IRS will not meet the original OMB and Treasury Department requirements.

## The DAR and DIU components are important parts of the DLP solution

The IRS deployed the DIM component and expanded its criteria, which is an important accomplishment; however, the DAR and DIU components are also important parts of the full solution. The DAR component provides the capability to scan data residing in data repositories to identify data vulnerable to exfiltration, at which point actions can be taken to address them. These data repositories can include data on workstations, server drives, or network shares. Once data at risk are identified, appropriate actions can be taken, including encrypting the data or deleting the data if they are not needed. This capability is especially useful given the large number of data repositories maintained by the IRS.

The purpose of the DIU component is to provide protection at the 'endpoint' (*i.e.*, workstation or laptop) as the proposed workflow for DIU allows users to correct their content before it is sent out. This is accomplished by monitoring and blocking confidential data from being printed, faxed, or copied to USB drives or other removable media.

The effectiveness of the two components is further increased when they are integrated into the overall DLP solution to work in concert with each other and with the DIM component. This includes using a common dashboard and reporting methodology for overall analytics and control and leveraging existing DIM policy sets for use as criteria. While the DIM component addresses the most obvious method of data exfiltration, PII is still at risk to insider threats by other avenues until the DAR and DIU components are deployed.

## Implementation progress of the DAR and DIU components was severely limited

The reasons for DLP project delays are varied, but they have been significant and ongoing, especially in regard to the development and deployment of the DAR and DIU components. The delays on the project as a whole and to the DAR and DIU components in particular have resulted in multiple changes to target release dates since the project began.

In TIGTA's September 2014 audit report on the IRS's progress with the implementation of the DLP solution, we reported that:

> *Based on its new projected implementation date of December 31, 2014, the IRS will have taken more than four years to build and develop its DLP solution.*

We concluded that:

> *Because of the length of time taken, TIGTA believes that stronger management oversight is needed to ensure that the DLP solution meets its new implementation date within budget.*[11]

On May 22, 2015, the IRS Cybersecurity office placed the DIM component into production, more than two years after the initial target release date and almost five months after its revised target date. In the years following the DIM component deployment, the project team continued to make incremental improvements to it. For example, during Calendar Year 2016, the DLP team expanded DIM criteria by including different Taxpayer Identification Number types in the SSN policy set, and in Calendar Year 2017, evaluation and testing of potential new policies (including credit card numbers) were performed. Also, new exclusions were added to the existing policies, including exclusions for specific websites that were producing false-positives. In August 2018, the DLP team deployed into production the DIM policy to detect and block transmission of credit card numbers and magnetic strip data. In addition, since the DIM component was deployed, the IRS has performed various activities to support the DLP solution as a whole, such as periodically upgrading the DLP software to successive versions and making upgrades to infrastructure.

The IRS has devoted resources to improving the DIM component since its deployment about four years ago, but the DAR and DIU components are still not operational. We identified the following factors that affected the SPIIDE Project's ability to deploy the DAR and DIU components of the DLP:

- **Efforts focused primarily on the DIM component implementation**

  According to the IRS, this was the primary cause of the overall project delay. The significant work required to deploy the DIM component and the post-implementation technical efforts encountered were key contributing factors delaying the timely deployment of the DAR and DIU components. After the DIM component was placed into production, the amount of work required to maintain and expand its capabilities was more than anticipated. When the amount of work to develop other capabilities such as Sensitive Image Recognition[12] was also considered, the IRS chose to reevaluate the overall focus of the project, which resulted in further delays to the DAR and DIU components.

  From May 2015 to November 2017, we found very little evidence where notable progress was made towards deployment of the other two components. Specifically, work related to the DAR and DIU components was sporadic and limited to various planning and

[11] TIGTA, Ref. No. 2014-20-087, *While the Data Loss Prevention Solution Is Being Developed, Stronger Oversight and Process Enhancements Are Needed for Timely Implementation Within Budget* p. 3 (Sept. 2014).
[12] An add-on capability to the DLP software that enables detection of sensitive text embedded in images.

testing activities as well as hardware acquisition through the end of Calendar Year 2017, when our audit started. The level of activity related to DAR and DIU development increased after that time, but as of the end of Calendar Year 2018, both components were still not ready for deployment. The original project charter had an estimated date of December 31, 2014, for full deployment of the DLP solution. The new estimated date based on the information for DAR and DIU deployment in the latest draft Work Breakdown Structure[13] is June 2020. This document also included planned start dates of September 2017 and December 2017 for DIU/Sensitive Image Recognition development and DAR development, respectively.

In November 2018, the IRS stated that additional changes were being made to the DLP implementation strategy, which would further affect the implementation dates for the DAR and DIU components. As of June 24, 2019, the IRS had not responded to our June 19, 2019, request for a DLP implementation strategy update.

- **Project management documentation was not always prepared or updated**

  Project documentation was not always prepared or updated as required after deployment of the DIM component, indicating inconsistent project management related to the DAR and DIU components. Documentation is an important part of project management, and various documents are required to be prepared and maintained during the course of the project. For example, Internal Revenue Manual 2.16.1, *Enterprise Life Cycle, Enterprise Life Cycle Guidance*,[14] lists specific documents required in order for projects to move through the stages of development. The project began in Calendar Year 2010, and we identified some required documentation that was approved in Calendar Year 2011, *e.g.*, the Project Charter and the Project Management Plan.[15]

  However, after the DIM was placed into production in May 2015, some required documents were still in draft form or had not been updated as required. For example, the Project Management Plan was originally approved in August 2011; however, even though it is a key project planning document, the plan had not been updated as required since the original approval. While the project team did continue to develop the DIM component over time, the documentation issues observed after it was deployed showed that working on the DAR and DIU components was a not a priority until Calendar Year 2018, when the project focus was reevaluated.

[13] A deliverable-oriented grouping of project elements that organizes and defines the total scope of a project. This project schedule is used to manage the tasks, task relationships, and resources needed to meet project goals.
[14] July 10, 2017.
[15] This document defines the project's scope of work and its approach to managing all project activities. Its purpose is to provide a framework for managing project activities and for completing the project successfully.

- **National Treasury Employees Union (NTEU) negotiations**

  In certain circumstances, the IRS is required to negotiate and reach a formal agreement with the NTEU prior to certain actions being taken. The IRS negotiated with the NTEU in regards to the DLP solution, and a Memorandum of Understanding was approved in July 2014 that set out certain stipulations and limitations related to how the DLP DIM solution affected bargaining unit employees. In the prior TIGTA audit, these negotiations were reported as having adversely affected project time frames, with information from IRS management indicating that the negotiations had taken a year. In September 2015, the IRS and NTEU signed an addendum to the original Memorandum of Understanding pertaining to the DIM component only. During this audit, the IRS stated that a meeting was scheduled with the NTEU to negotiate a Memorandum of Understanding related to the implementation of the DAR and DIU components. IRS management has again cited the negotiations as the cause of delays with project implementation.

## *Further project delays to the DAR and DIU components could result in additional inefficient use of resources*

Executive Order 13589 (November 9, 2011), *Promoting Efficient Spending*, requires Federal agencies to establish controls to ensure that they are not paying for unused or underutilized information technology equipment, software, or services. In addition, IRS policy states that information technology governance is a function of internal control within the IRS, and the primary objective of governance is to ensure that assigned investment, program, and project objectives are met; risks are managed; and expenditures are sound. Accordingly, IRS management is directly responsible for ensuring that funds allocated to information technology projects, such as the SPIIDE Project's DLP program, are not being misused or wasted.

As part of the SPIIDE Project, the IRS originally contracted for the DLP solution software in early Calendar Year 2011. At that time, the IRS purchased 110,000 licenses when its workforce was about 104,000 employees (extra licenses were needed for contractors), and another 30,000 licenses for the Treasury Department to use, for a total of 140,000 licenses. However, the IRS workforce subsequently decreased, and by Calendar Year 2015, the total number of employees was substantially smaller (about 90,000). Recognizing this reduction, in Calendar Year 2015, the Treasury Department took over the administration of the DLP license renewal contract, including the DIM, DAR, and DIU components, and transitioned to a Departmentwide contract for the same 140,000 licenses, all of which were then made available for use to all Treasury Department bureaus.

The IRS remained the largest user of the licenses. Because of the Treasury Department's actions, the IRS stopped contracting directly for the DLP license renewals and began acquiring

them using an Interagency Agreement[16] with the Treasury Department under its Franchise Fund Shared Services Program.[17]

Figure 3 shows the full DLP cost based on the Treasury Department's contract amounts for Fiscal Years 2016 through 2019 and the associated amounts attributable to the DAR and DIU components.

### *Figure 3: DLP Contract Amounts for Fiscal Years 2016 Through 2019*

| | Fiscal Year | | | | |
|---|---|---|---|---|---|
| | **2016** | **2017** | **2018** | **2019** | **Total** |
| **Treasury Department Cost for DLP (full agreed contract price per fiscal year)** | **$565,100** | **$625,900** | **$654,500** | **$692,700** | **$2,538,200** |
| **DIM and Remote Assistance/ Technical Support Portion (per contract)** | $211,000 | $253,000 | $277,900 | $294,100 | $1,036,000 |
| **DAR Portion (per contract)** | $175,900 | $161,900 | $137,900 | $146,000 | $621,700 |
| **DIU Portion (per contract)** | $178,300 | $211,000 | $238,700 | $252,600 | $880,600 |
| **Total Cost of DAR and DIU (per contract)** | **$354,200** | **$372,900** | **$376,600** | **$398,600** | **$1,502,300** |

*Source: TIGTA's analysis of the DLP licenses. Some totals may not compute due to rounding.*

The Treasury Department uses the Treasury Franchise Fund to assign and allocate shared costs to the requesting bureaus. The Treasury Department then bills the bureaus monthly for their share of the services. In Fiscal Year 2019, the Treasury Franchise Fund estimated that the overall DLP cost was $692,700, and the IRS share was set at about 80 percent of that amount.

We obtained and analyzed the contract documents pertinent to the DLP solution software license costs, which provided separate license information for the three individual components. We limited our analysis to the costs incurred for the three components from Fiscal Years 2016 through 2019. We used the IRS's Fiscal Year 2019 Treasury Franchise Fund share (*i.e.*, about 80 percent) to determine the cost allocated to the IRS for each component for prior fiscal years.

The significant delays with the deployment of the DAR and DIU components of the DLP solution resulted in the inability to use the capabilities associated with these two components.

[16] A written agreement entered into between two Federal agencies, or major organizational units within an agency, that specifies the goods to be furnished or tasks to be accomplished by one agency (the service agency) in support of the other (the requesting agency).
[17] The Shared Services Program with the Treasury Franchise Fund provides common administrative services that benefit customers both within the Treasury Department and outside agencies.

However, the IRS continued to pay for DLP license renewal costs for these two components per the terms of the Interagency Agreement.

Figure 4 shows the total amount paid for unused software based on the total contract costs associated with the DAR and DIU components and the cost allocation for the IRS per the Interagency Agreement terms.

### *Figure 4: Analysis of Amounts Paid for the DLP Components That Were Not Used by the IRS From Fiscal Years 2016 Through 2019*

| | Fiscal Year | | | | |
|---|---|---|---|---|---|
| | **2016** | **2017** | **2018** | **2019** | **Total** |
| **Total of DAR and DIU Portions (per contract from Figure 3)** | **$354,200** | **$372,900** | **$376,600** | **$398,600** | **$1,502,300** |
| **DAR Amount (IRS portion)** | $140,800 | $129,600 | $110,400 | $116,900 | $497,700 |
| **DIU Amount (IRS portion)** | $142,700 | $168,900 | $191,100 | $202,300 | $705,000 |
| **Total Amount Paid for Unused Software by IRS** | **$283,500** | **$298,500** | **$301,500** | **$319,200** | **$1,202,700** |

*Source: TIGTA's analysis of the DLP licenses.*

During this period, a cost of about $1.5 million was incurred by the Treasury Department for license renewals associated with the DAR and DIU components. The IRS was responsible for about 80 percent of this cost based on the terms of the Interagency Agreement with the Treasury Department. Therefore, the IRS was responsible for paying approximately $1.2 million for software that was not deployed into production, *i.e.*, not in use, over a four-year period. We did not include the costs incurred under these contracts for remote assistance/technical support for the DLP solution that was allocated to the IRS.

When the IRS initially contracted for the DLP software in Fiscal Year 2011, all three components were expected to be deployed by the end of Calendar Year 2014. By the time the Treasury Department took over administration of the contract in February 2015, the original projected release dates had elapsed and none of the three components had been deployed. From that point, delays continued to affect the project, resulting in the DAR and DIU components not being deployed as originally planned. The IRS estimates that both components will be implemented by June 15, 2021.

To achieve the full functionality envisioned for the DLP solution, all three components must be deployed into a production environment. Therefore, the delays related to the DAR and DIU implementation are preventing full compliance with the OMB and other requirements and the realization of the full project benefits, including the protection of PII and the efficient use of resources.

## Recommendations

The Chief Information Officer should:

Recommendation 1**:** Deploy the DAR and DIU components of the DLP solution.

> ***Management's Response:*** The IRS agreed with this recommendation. The Cybersecurity office will deploy the DAR and DIU components.

***Recommendation 2:*** Ensure that project documents are prepared and maintained as required for effective project management, which should help ensure the successful delivery of the final two components of the DLP solution.

> ***Management's Response:*** The IRS agreed with this recommendation. The SPIIDE Project team will ensure that project documents are consistently prepared and maintained during the deployment and delivery of the DAR and DIU components.

***Recommendation 3:*** Ensure that any issues requiring negotiations with the NTEU related to the SPIIDE Project are identified and negotiations started promptly to reduce potential adverse impacts on project timelines.

> ***Management's Response:*** The IRS agreed with this recommendation. The Memorandum of Understanding is currently in the process of concurrence signatures. The SPIIDE Project team will notify the NTEU of any issues as stipulated in the Memorandum of Understanding agreement regarding the production implementation of the DIU and DAR components.

## Appendix I

# Detailed Objective, Scope, and Methodology

The overall objective of this review was to evaluate whether the IRS has properly implemented controls to prevent data loss, including data exfiltration of personal information. To accomplish our objective, we:

I. Determined the overall status of the DLP project, whether it was effectively managed to meet planned milestones and minimize project costs, and if the criteria used was in accordance with relevant guidance.

   A. Determined when the solution was to be fully implemented and the causes of any delays.

   B. Obtained and analyzed contracts and financial documents to determine if the IRS effectively utilized the DLP licenses allocated to it per the contract.

   C. Determined whether the DLP policies and procedures were in accordance with applicable criteria.

II. Determined whether the current operational DLP component was effectively identifying and blocking the PII the IRS was trying to protect.

   A. Determined if the information protected by the DIM solution was consistent with the written scope of the project.

   B. Reviewed Government standards to determine what other types of PII could have been included in the DIM scope and determined if the IRS considered these items. We determined why some items were not included.

   C. Determined if the DIM component was functioning as intended to successfully identify and block relevant data, taking into account the intended IRS scope of data protection.

   D. Selected a judgmental sample[1] of 56 of the 1,561 potential incidents generated from May 1 through May 14, 2018, and determined if identified potential incidents were processed by following the correct procedures for routing and remediation.

### *Internal controls methodology*

Internal controls relate to management's plans, methods, and procedures used to meet their mission, goals, and objectives. Internal controls include the processes and procedures for

[1] A judgmental sample is a nonprobability sample, the results of which cannot be used to project to the population.

planning, organizing, directing, and controlling program operations. They include the systems for measuring, reporting, and monitoring program performance. We determined that the following internal controls were relevant to our audit objective: OMB memoranda, Internal Revenue Manual sections, National Institute of Standards and Technology and Treasury Department guidelines, and other procedures related to implementing the DLP solution to monitor PII. We evaluated these controls by interviewing IRS management and staff and reviewing relevant documentation from the National Institute of Standards and Technology, the OMB, the Treasury Department, and the IRS. We also reviewed other relevant supporting documentation, such as DLP incident reports and documents supporting the procurement of the DLP solution.

## Appendix II

# Major Contributors to This Report

Danny Verneuille, Assistant Inspector General for Audit (Security and Information Technology Services)
Kent Sagara, Director
Jason McKnight, Acting Audit Manager
Ryan Perry, Acting Audit Manager
Steven Stephens, Lead Auditor
Midori Ohno, Senior Auditor
Linda Nethery, Information Technology Specialist

## Appendix III

# Report Distribution List

Deputy Commissioner for Operations Support
Chief Information Officer
Deputy Chief Information Officer for Operations
Associate Chief Information Officer, Cybersecurity
Director, Cybersecurity Architecture and Implementation
Director, Cybersecurity Operations
Director, Enterprise Audit Management

## Appendix IV

# Outcome Measure

This appendix presents detailed information on the measurable impact that our recommended corrective actions will have on tax administration. This benefit will be incorporated into our Semiannual Report to Congress.

### Type and Value of Outcome Measure:

- Inefficient Use of Resources – Potential; $1.2 million (see page 9).

### Methodology Used to Measure the Reported Benefit:

The IRS pays for the use of the DLP software through an Interagency Agreement with the Treasury Department, which contracts with the vendor for the software. The terms of the Interagency Agreement dictate that the IRS is responsible for a percentage of the contract cost, as determined by the Treasury Department. For Fiscal Year 2019, the Treasury Department set the IRS's share of the DLP cost at 80.06 percent.

From Fiscal Years 2016 through 2019, the Treasury Department paid $2,538,190 for the entire DLP solution. Our analysis of the contracts found that $1,502,253 was attributed to the DIU and the DAR components of the DLP solution. By applying the IRS's share of the contract costs of 80.06 percent, we calculated that the IRS paid $1,202,704 for the DIU and the DAR components not deployed into production during the four-year period.

## Appendix V

## *Management's Response to the Draft Report*

DEPARTMENT OF THE TREASURY
INTERNAL REVENUE SERVICE
WASHINGTON, D.C. 20224

CHIEF INFORMATION OFFICER

July 31, 2019

MEMORANDUM FOR MICHEAL E. MCKENNEY,
DEPUTY INSPECTOR GENERAL FOR AUDIT

FROM: Nancy A. Sieger
Acting Chief Information Officer

SUBJECT: Draft Audit Report – The First Phase of the Data Loss Prevention Is Working As Intended, but the Remaining Phases Continue to Experience Delays (Audit # 201820003) (e-trak #2019-13801)

Thank you for the opportunity to review the draft audit report and meet with the audit team to discuss report observations. We appreciate your report's acknowledgement that the Data-in-Motion (DIM) component of the Data Loss Prevention (DLP) solution has been implemented and that the Internal Revenue Service (IRS) has continued to expand its functionality to enhance the protection of Personally Identifiable Information (PII). Additionally, we appreciate your acknowledgement that the DIM solution is working as intended and that potential incidents were reviewed and resolved properly across the numerous organizations and layers of management engaged in the successful execution of this program.

The deployment of the full DLP product suite is just one of numerous ongoing efforts to secure our systems and protect sensitive information. Our dedicated focus on Cybersecurity has positioned the IRS to withstand approximately 1.4 billion cyberattacks annually (including denial-of-service attacks, unsuccessful intrusion attempts, probes or scans, and other unauthorized connectivity attempts). Many of these attempts are sophisticated in nature or represent advanced, persistent threats. To continue successfully defending our systems and combatting tomorrow's threats, the IRS is committed to continued investments and program improvements in our cyber defenses.

The IRS is committed to the protection of all data from intentional or unintentional disclosures, including PII entrusted to us by the taxpayers we serve. To that end, the IRS has taken significant steps to enhance our protections and defenses against malicious acts and expand our use of continuous application security monitoring to detect and prevent potential incidents. These improvements include implementation of the first phase of the Department of Homeland Security (DHS) sponsored Continuous Diagnostics and Mitigation (CDM) program, which required significant resources, personnel, and management attention from the organization also responsible for deploying the DLP solution.

While we recognize that continued prioritization of limited resources is an ongoing reality, we remain committed to continue our progress toward full implementation of the DLP solution. Additionally, we agree with your report that implementation challenges occurred with deployment of the Data-in-Use (DIU) and Data-at-Rest (DAR) components of the DLP solution. However, we do not support the conclusion that these deployment delays resulted in an inefficient use of resources.

In evaluating complex acquisition of software solutions, it is extremely important to evaluate the Total Cost of Ownership (TCO). Large government agencies, such as the IRS, often obtain heavily discounted pricing that results in ownership of a full product suite for far less than the cost to own one specific component. In fact, in some cases opting to purchase only one component can result in the government paying a higher total price for the acquisition. On this basis, the IRS supports the Treasury departmental decision to consider the TCO in its acquisition strategy.

In 2015, the Treasury Department awarded a multi-year Firm Fixed Price (FFP) contract for DLP software based on the products and licensing requirements of all bureaus. The terms of that contract contained extraordinary price reductions that afforded all Treasury Department bureaus the ability to test and progress toward implementation of the full suite of DLP software for 90 percent less than the General Services Administration (GSA) price for owning the single DLP DIM component that was effectively deployed by the IRS. Due to these significant price reductions, the IRS's holistic view of the TCO for the acquisition reflects that the pricing structure and license sharing across all Treasury bureaus was extremely advantageous to the government. The total savings exceeded nearly $10 million, which far surpassed the $1.2 million for the four-year period noted in your report. In the light of the total savings, our view is that the overall contractual cost avoidance and the planned implementation of the DIU and DAR functions are not an inefficient use of resources.

Attached is our detailed corrective action plan to implement the audit report's recommendations. The IRS values your continued support and the assistance your organization provides. If you have any questions, please contact me at (202) 317-5000 or Jamie Plummer, Acting Senior Manager of Program Oversight Coordination, at (240) 613-2191.

Attachment

Attachment

Draft Audit Report The First Phrase of the Data Loss Prevention Solution Is Working As Intended, but the Remaining Phases Continue to Experience Delays (Audit # 201820003)

**RECOMMENDATION #1:** The Chief Information Officer should deploy the DAR and DIU components of the DLP solution.

**CORRECTIVE ACTION #1**: The Internal Revenue Service (IRS) agrees with this recommendation. Cybersecurity will deploy Data Loss Prevention (DAR) and Data-in-Use (DIU) components.

**IMPLEMENTATION DATE :** 6/15/2021

**RESPONSIBLE OFFICIALS:** Associate Chief Information Officer, Cybersecurity

**CORRECTIVE ACTION MONITORING PLAN:** We enter accepted Corrective Actions into the Joint Audit Management Enterprise System (JAMES) and monitor them on a monthly basis until completion.

**RECOMMENDATION #2:** The Chief Information Officer should ensure project documents are consistently prepared and maintained as required for effective project management, which should ensure the successful delivery of the final two components of the DLP solution.

**CORRECTIVE ACTION #2**: The Internal Revenue Service (IRS) agrees with this recommendation. The Cybersecurity Safeguarding Personally Identifiable Information Data Extract (SPIIDE) project team will ensure project documents are consistently prepared and maintained, during the deployment and delivery of the Data-at-Rest (DAR) and Data-in-Use (DIU) components.

**IMPLEMENTATION DATE:** 10/15/2021

**RESPONSIBLE OFFICIALS:** Associate Chief Information Officer, Cybersecurity

**CORRECTIVE ACTION MONITORING PLAN:** We enter accepted Corrective Actions into the Joint Audit Management Enterprise System (JAMES) and monitor them on a monthly basis until completion.

**RECOMMENDATION #3:** The Chief Information Officer should ensure that any issues requiring negotiations with the NTEU related to the SPIIDE Project are identified and negotiations started promptly to reduce potential adverse impacts on project timelines.

**CORRECTIVE ACTION #3**: The Internal Revenue Service (IRS) agrees with this recommendation. The National Treasury Employees Union Memorandum of Understanding (NTEU MOU) is currently in the process of concurrence signatures. The Safeguarding Personally Identifiable Information Data Extracts (SPIIDE) Project team

Attachment

Draft Audit Report The First Phrase of the Data Loss Prevention Solution Is Working As Intended, but the Remaining Phases Continue to Experience Delays (Audit # 201820003)

will notify NTEU of any issues as stipulated in the MOU agreement regarding the production implementation of Data-In-Use (DIU) and Data-At-Rest (DAR).

**IMPLEMENTATION DATE:** 6/15/2020

**RESPONSIBLE OFFICIALS:** Associate Chief Information Officer, Cybersecurity

**CORRECTIVE ACTION MONITORING PLAN:** We enter accepted Corrective Actions into the Joint Audit Management Enterprise System (JAMES) and monitor them on a monthly basis until completion.

## Appendix VI

# Office of Audit Comments on Management's Response

While IRS management agreed with all of the recommendations in the report, the IRS disagreed that deployment delays for two DLP components resulted in an inefficient use of resources. In its management response to the draft report, the IRS asserted that the significant price reduction below the General Services Administration price for the DLP software, including all three components, far surpassed the money spent on the DIU and DAR components that were not implemented as expected.

**Management's Response:** *In 2015, the Treasury Department awarded a multi-year Firm Fixed Price contract for DLP software based on the products and licensing requirements of all bureaus. The terms of that contract contained extraordinary price reductions that afforded all Treasury Department bureaus the ability to test and progress toward implementation of the full suite of DLP software for 90 percent less than the General Services Administration price for owning the single OLP DIM component that was effectively deployed by the IRS. Due to these significant price reductions, the IRS's holistic view of the Total Cost Ownership for the acquisition reflects that the pricing structure and license sharing across all Treasury bureaus was extremely advantageous to the government. The total savings exceeded nearly $10 million, which far surpassed the $1.2 million for the four-year period noted in your report. In the light of the total savings, our view is that the overall contractual cost avoidance and the planned implementation of the DIU and DAR functions are not an inefficient use of resources.*

> **Office of Audit Comment:** While the Department of the Treasury obtained the extraordinary price reductions for the DLP solution, the IRS inefficiently used its resources when it did not implement two components of the DLP solution and make full use of the purchased software capabilities. The IRS paid $1.2 million from Fiscal Years 2016 through 2019 for unused licenses for the DIU and DAR components of the DLP solution.