



*The Return Review Program Increases
Fraud Detection; However, Full Retirement
of the Electronic Fraud Detection System
Will Be Delayed*

September 25, 2017

Reference Number: 2017-20-080

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.

Redaction Legend:

2 = Law Enforcement Techniques/ Procedures and Guidelines for Law Enforcement Investigations or Prosecutions

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.



HIGHLIGHTS

THE RETURN REVIEW PROGRAM INCREASES FRAUD DETECTION; HOWEVER, FULL RETIREMENT OF THE ELECTRONIC FRAUD DETECTION SYSTEM WILL BE DELAYED

Highlights

Final Report issued on September 25, 2017

Highlights of Reference Number: 2017-20-080 to the Internal Revenue Service Chief Information Officer.

IMPACT ON TAXPAYERS

The goal of implementing the Return Review Program (RRP) is to replace the older Electronic Fraud Detection System (EFDS) with an automated system that would enhance the IRS's capabilities to prevent, detect, and resolve criminal and civil noncompliance. The IRS developed the RRP with the end goal of having one fraud detection system for systemic case selection that identified more fraudulent returns at a lower false detection rate than the EFDS and the Dependent Database, resulting in reduced taxpayer burden.

WHY TIGTA DID THE AUDIT

The IRS reported that the long-term limitations of the EFDS include its inability to keep pace with increasing levels of fraud or to serve the IRS's evolving compliance needs. The IRS has implemented incremental RRP functionality since launching its first pilot of Identity Theft models in April 2014 to identify how many more electronically filed tax returns the RRP would have selected into inventory versus the EFDS. The overall objective of this review was to determine if the RRP system can identify all fraud currently identified by other existing fraud detection systems and to assess the EFDS retirement plans.

WHAT TIGTA FOUND

The IRS retired the EFDS Identity Theft models for the 2016 Filing Season and the Non-Identity Theft models for the 2017 Filing Season after

demonstrating that the RRP met the IRS's objectives of identifying more fraudulent returns at a lower false detection rate. The IRS leveraged the RRP to detect some of the identity theft returns selected by the Dependent Database. The IRS also significantly expanded the RRP's systemic use of ***2*****new data elements in the 2017 Filing Season to detect identity theft tax returns.

The IRS addressed a prior TIGTA recommendation by developing a plan to retire the EFDS. In December 2015, the IRS Executive Steering Committee unanimously approved the EFDS Retirement Strategy, which called for the retirement of 19 components by December 2018. Six of the eight components with a retirement date of December 2015 or December 2016 were retired timely. The other two components are now scheduled to be retired in October 2017. The remaining 11 of 19 EFDS components are related to Enterprise Case Management and have retirement dates in December 2017 or December 2018. The IRS initiated a separate Enterprise Case Management project in January 2015, to be implemented in December 2018. However, in February 2017, the IRS suspended the Enterprise Case Management project due to insufficient funding and staffing. In addition, the IRS determined that the software product selected for Enterprise Case Management cannot support an enterprise-wide deployment.

WHAT TIGTA RECOMMENDED

TIGTA made no recommendations in the report. However, key IRS officials reviewed this report prior to its issuance and agreed with the facts and conclusions presented.



TREASURY INSPECTOR GENERAL
FOR TAX ADMINISTRATION

DEPARTMENT OF THE TREASURY
WASHINGTON, D.C. 20220

September 25, 2017

MEMORANDUM FOR CHIEF INFORMATION OFFICER

FROM: Michael E. McKenney
Deputy Inspector General for Audit

SUBJECT: Final Audit Report – The Return Review Program Increases Fraud Detection; However, Full Retirement of the Electronic Fraud Detection System Will Be Delayed (Audit # 201620019)

This report presents the results of our review to determine if the Return Review Program system can identify all fraud currently identified by other existing fraud detection systems and to assess the Electronic Fraud Detection System retirement plans. This audit is included in our Fiscal Year 2017 Annual Audit Plan and addresses the major management challenge of Improving Tax Compliance.

We made no recommendations as a result of the work performed during this review. However, key Internal Revenue Service officials reviewed this report prior to its issuance and agreed with the facts and conclusions presented.

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



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Table of Contents

<u>Background</u>	Page 1
<u>Results of Review</u>	Page 3
<u>The Return Review Program Is Better Positioned to Meet Business Objectives and Address the Changing Nature of Identity Theft</u>	Page 3
<u>Results From Recent Filing Seasons Support the Decision to Retire the Electronic Fraud Detection System Models</u>	Page 5
<u>The Scope of the Return Review Program Was Expanded to Include Some Dependent Database Identity Theft Detection Functionality</u>	Page 8
<u>The Systemic Utilization of Data Elements in the Return Review Program Was Expanded in the 2017 Filing Season</u>	Page 9
<u>Progress Was Made Towards Retiring the Electronic Fraud Detection System, but It Will Not Meet the Planned System Retirement Date of December 2018</u>	Page 13
 Appendices	
<u>Appendix I – Detailed Objective, Scope, and Methodology</u>	Page 16
<u>Appendix II – Major Contributors to This Report</u>	Page 18
<u>Appendix III – Report Distribution List</u>	Page 19
<u>Appendix IV – Glossary of Terms</u>	Page 20



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Abbreviations

CRISP-DM	Cross-Industry Standard Process for Data Mining
DDb	Dependent Database
ECM	Enterprise Case Management
EFDS	Electronic Fraud Detection System
FDR	False Detection Rate
IDT	Identity Theft
IRS	Internal Revenue Service
Non-IDT	Non-Identity Theft
RRP	Return Review Program
TIGTA	Treasury Inspector General for Tax Administration



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Background

The Internal Revenue Service (IRS) implemented the Electronic Fraud Detection System (EFDS)¹ in 1994 to identify questionable and potentially fraudulent tax returns. Over time, the IRS determined that numerous inefficiencies and operational challenges rendered the EFDS too risky to maintain, upgrade, or operate long term. The IRS reported that the long-term limitations of the EFDS include its inability to keep pace with increasing levels of fraud or to serve the IRS's evolving compliance needs. In 2009, the IRS began a project to replace the EFDS with the Return Review Program (RRP), an automated system that would enhance the IRS's capabilities to prevent, detect, and resolve criminal and civil noncompliance.

The IRS developed the RRP with the end goal of having one fraud detection system of record for systemic case selection that identified more fraudulent tax returns at a lower False Detection Rate.

The IRS developed the RRP with the end goal of having one fraud detection system for systemic case selection that identified more fraudulent tax returns at a lower False Detection Rate (FDR)² than the EFDS and the Dependent Database (DDb),³ resulting in reduced taxpayer burden. To achieve the goal, the IRS developed a release plan to deliver incremental RRP functionality over multiple filing seasons. However, RRP development was suspended in January 2014 to allow the IRS time to evaluate the performance and design of the first RRP release and to revisit strategic business fraud detection goals. To exit the suspension, the IRS developed a restart plan that was approved by the Executive Steering Committee in January 2015. In addition, the IRS realigned the RRP to be deployed as the Enterprise Anomaly Detection System, with case management functionality to be delivered as a separate Enterprise Case Management (ECM) project.

In April 2014, within 60 days of deploying the RRP into production, the IRS launched what is referred to as the "RRP Identity Theft (IDT) Pilot." In the pilot, the RRP made IDT selections⁴ one day a week to identify potential fraudulent returns. The returns selected by the RRP pilot were added into the EFDS reporting database in an effort to identify additional potential fraud

¹ See Appendix IV for a glossary of terms.

² The rate at which a test result incorrectly indicates that a particular condition or attribute is present. Some of the IRS documentation refers to FDRs, while other IRS documentation refers to False Positive Rates. The two terms are synonymous; however, FDR is the measurement now used by the IRS.

³ A risk-based audit selection tool used by the IRS to identify tax returns for audit. It is made up of a collection of information databases that include birth certificate information and court documents used to establish a relationship and residency between a taxpayer and the qualifying children claimed on a tax return.

⁴ Questionable tax returns identified by fraud detection systems.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

previously undetected by the EFDS and the DDb. The goal of the pilot was to expand IDT detection capabilities at an FDR lower than the legacy fraud detection systems.

The IRS reported that the pilot was successful. Specifically, the RRP increased fraud detection – 59.4 percent over the IDT found by the EFDS or the DDb. In addition, the RRP had the lowest FDR (6 percent) compared to the EFDS (7.5 percent) and the DDb (19.7 percent).

Based on the reported success of the RRP IDT pilot in the 2014 Filing Season, the IRS started using the RRP IDT models to select electronically filed tax returns in the 2015 Filing Season. Also in the 2015 Filing Season, the IRS piloted the RRP Non-Identity Theft (Non-IDT) fraud models on electronically filed tax returns. Beginning in the 2016 Filing Season, the RRP was the only system capable of making both IDT and Non-IDT selections for electronically filed tax returns. Prior to the start of the 2017 Filing Season, the IRS decided to retire the EFDS Non-IDT models, making the 2017 Filing Season the first year in which the EFDS did not run models in parallel with the RRP. See Figure 1 for the operational status of the EFDS and RRP fraud detection models for the 2014 through 2017 Filing Seasons.

**Figure 1: Operational Status of the EFDS and RRP
Fraud Detection Models for the 2014 Through 2017 Filing Seasons**

Filing Season	IDT Models		Non-IDT Models	
	EFDS	RRP	EFDS	RRP
2014	Operational	Pilot – April 2014	Operational	
2015	Operational	Operational	Operational	Pilot – September 2015
2016	Retired	Operational	Operational	Operational
2017		Operational	Retired	Operational

Source: IRS – Timeline of Return Review Program Expansion in Fraud Detection provided data.

This review was performed in the Information Technology organization’s Enterprise Program Management Office and Applications Development function and the Wage and Investment Division’s Return Integrity and Compliance Services function in the New Carrollton Federal Building in Lanham, Maryland, during the period August 2016 through July 2017. 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. Detailed information on our audit objective, scope, and methodology is presented in Appendix I. Major contributors to the report are listed in Appendix II.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Results of Review

The Return Review Program Is Better Positioned to Meet Business Objectives and Address the Changing Nature of Identity Theft

In a December 2015 report, the Treasury Inspector General for Tax Administration (TIGTA) reported⁵ that the RRP is not identifying all of the IDT tax returns identified by other systems. TIGTA recommended that the Commissioner, Wage and Investment Division, ensure that IDT tax returns identified by the DDb or the EFDS are selected by the RRP prior to replacing the IDT capabilities of these systems. The IRS agreed that retirement of existing systems should not compromise or reduce its ability to detect potentially fraudulent tax returns. The IRS also indicated that it already had a plan which it believed adequately addressed the intention of the recommendation.

We reviewed the IRS's plan to address the December 2015 TIGTA recommendation. The IRS provided us with a document titled *EFDS to RRP Traceability* completed in August 2015. The *EFDS to RRP Traceability* documented the current state of the EFDS and traced existing EFDS functionality to corresponding elements in the RRP. From this exercise, the IRS documented 14 EFDS capability statements, 143 EFDS functionality statements, and 76 EFDS requirement statements. The IRS Requirements Engineering Program Office provided the following definitions to explain the difference between a capability, functionality, and requirement:

- Capability relates directly to the project scope and often is the highest level of requirement.
- Functionality specifies what the information system is expected to do.
- Requirement represents a condition or capability needed by an end user to solve a problem or achieve an objective. A requirement should be used as the basis for design and development.

Subsequently, the IRS traced the 143 functionality statements to the corresponding functionality in the RRP. The outcomes of the traceability exercise were grouped into three categories:

- **Gap** – represents EFDS functionality that is not documented as existing or future RRP functionality.

⁵ TIGTA, Ref. No. 2016-40-008, *Continued Refinement of the Return Review Program Identity Theft Detection Models Is Needed to Increase Detection* (Dec. 2015).



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

- **Watch List Item** – represents EFDS functionality that has not been fully defined in RRP requirements and needs to be further analyzed or developed as part of either RRP or ECM activities.
- **Opportunity** – represents functionality that is not present in either the EFDS or RRP but may be beneficial or business critical and therefore should be considered in the future state solution.

The traceability work resulted in the identification of six gaps, 33 EFDS watch list items, and six opportunities. The gaps and EFDS watch list items represented existing functionality required to retire the EFDS.

We analyzed the *EFDS to RRP Traceability* at the EFDS capability and functionality statement levels. We concluded that all 14 EFDS capability statements traced to one or more of the 23 RRP capability statements. In addition, all 143 EFDS functionality statements traced to one or more of the 23 RRP capability statements. In conclusion, all of the EFDS capabilities, functionalities, and requirements that the IRS cross-walked either showed traceability to one or more of the 23 RRP capability statements or, as warranted, a gap, an EFDS watch list item, or an opportunity

During the audit, the IRS provided updated information to reflect the status of the gaps, EFDS watch list items, and opportunities related to the manual identification of fraud that the IRS viewed as essential because identification could not be done systemically at this time. However, during meetings with the IRS to review our analysis, we learned that the *EFDS to RRP Traceability*, and the associated gap analysis, did not have any impact on the RRP's ability to systemically detect or prevent fraud. As a result, we did not pursue our analysis of the *EFDS to RRP Traceability* any further.

In discussions with Applications Development function personnel, they stated that despite the IRS Wage and Investment Division's agreement with TIGTA's prior recommendation, some tax returns selected by the EFDS and/or the DDb are not selected by the RRP. The reason is that no two fraud detection systems are going to find exactly the same cases of fraud when deploying predictive models developed using different proprietary machine learning algorithms as its primary detection method. Further, the Applications Development function personnel repeatedly stated that the RRP is not designed to identify the same returns selected by the EFDS and DDb. The goal of the RRP is to ensure that: 1) there is no gap in the ability to develop models using the RRP's proprietary analytics software and 2) the RRP proprietary analytics software results in greater fraud detection at a lower FDR than the EFDS.

A key justification for the Application Development function's positioning of the RRP is that, since 2014, the trend in IDT has changed significantly. The IRS defines IDT as the use of a stolen Social Security Number to file a tax return claiming a fraudulent refund. In prior years, the identity thief submitted a fraudulent tax return using the legitimate taxpayer's Social Security Number, often without knowing any additional information required on the Federal tax return



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

*****2*****. In recent years, with the increasing frequency of data breaches, identity thieves have access to more information about taxpayers. A report from the Identity Theft Resource Center in January 2017 stated that U.S. companies and government agencies suffered a record 1,093 data breaches in 2016, a 40 percent increase from 2015. Notable examples include one technology company that had all its Form W-2, *Wage and Tax Statement*, documents for current and past employees stolen. Another example was the data breach of the IRS in which hackers accessed the information of more than 700,000 individual taxpayers. Because identity thieves have more data about taxpayers, they can input more legitimate tax return data on the fraudulent return, making it harder for IRS fraud detection systems to differentiate an IDT tax return from a legitimate tax return. Between 2014 and 2016, the percentage of IDT tax returns that have more legitimate tax return data grew from 8 percent of IDT returns to 50 percent.

We believe that the RRP is better positioned than the EFDS to address the changing nature of IDT due to the following two reasons.

- 1) The EFDS uses models to generate one fraud score for each return. In contrast, RRP models generate a set of predictive scores for every return. This enables the RRP to individually assess tax returns across all IDT and Non-IDT fraudulent categories.
- 2) The RRP has a more robust business rules engine compared to the EFDS. As such, it has greater flexibility to adjust to new emerging fraud trends.

The following section details how the RRP better meets the IRS objectives of delivering greater fraud detection at a lower FDR than the EFDS.

[Results From Recent Filing Seasons Support the Decision to Retire the Electronic Fraud Detection System Models](#)

The IRS uses the Cross-Industry Standard Process for Data Mining (CRISP-DM) to close gaps between the EFDS and the RRP. The CRISP-DM is a proven, well-structured process for predictive modeling consisting of the following six phases:

1. **Business Understanding** – focuses on understanding the project objectives and requirements.
2. **Data Understanding** – enables the IRS to enhance understanding of relevant data sources and identify any data quality problems.
3. **Data Preparation** – covers transformation, integration, and cleaning activities needed to prepare the data for modeling.
4. **Modeling** – applies various analytic techniques to develop predictive models.
5. **Evaluation** – focuses on thoroughly assessing, refining, and validating the model. In this phase, the IRS uses Receiver Operating Characteristic curves to quantitatively measure



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

performance. The curves measure the False Positive Rate of a predictive model for a given target fraud detection rate. The IRS runs each model through numerous iterations to test which combination of settings produce the best predictive fraud detection model.

6. **Deployment** – involves moving new models into production to score returns. This phase includes rigorous testing to validate that the models are correctly deployed.

The predictive model with the best results on the Receiver Operating Characteristic curve becomes the champion model that the RRP deploys to production and uses during the upcoming filing season to detect fraud. The IRS has used CRISP-DM to develop all the analytic models deployed in the EFDS and the RRP since 2011. By following the same process to develop models as the EFDS, the IRS believes it ensures consistent fraud detection between the EFDS and the RRP.

The IRS retired the EFDS IDT models for the 2016 Filing Season. The IRS believed the RRP was producing vastly greater results than the EFDS in the prevention of IDT. We obtained data from the IRS for the 2015 Filing Season to evaluate the IRS’s decision. The filing season data showed that collectively, the EFDS, the RRP, and the DDb identified a total of 700,560 confirmed IDT tax returns, representing \$3.92 billion in revenue. Some of the returns were only selected by one fraud detection system. Some of the IDT tax returns were selected by multiple systems (*e.g.*, selected by both the RRP and the EFDS or by the RRP and the DDb). Figure 2 summarizes the number of IDT returns selected by each individual fraud detection system as well as IDT returns selected by a combination of systems.

Figure 2: 2015 Filing Season Identity Theft Selections by System

Selection Source	Confirmed IDT	Revenue Protected	FDR
RRP Selections	256,204	\$1.88 Billion	28.6%
DDb Selections	128,605	\$408 Million	64.7%
EFDS Selections	9,416	\$60 Million	32.5%
Selected by Multiple Systems	306,335	\$1.58 Billion	8.6%
Total ⁶	700,560	\$3.92 Billion	

Source: IRS – *Timeline of Return Review Program Expansion in Fraud Detection* provided data.

As Figure 2 shows, the EFDS IDT model selections that were not selected by any other fraud detection system accounted for \$60 million, which is 1.5 percent of the total \$3.92 billion IRS

⁶ The term total refers to all systemic selections. There are additional tax returns selected by manual processes that are not reflected in Figure 2.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

IDT revenue protection. In comparison, the RRP IDT model selections that were not selected by any other fraud detection system totaled \$1.88 billion, accounting for 47.8 percent of the total \$3.92 billion IRS IDT revenue protection.

We asked the IRS if running the EFDS IDT models for another filing season would enable the RRP to close the fraud detection gap. The IRS responded that the RRP can identify the IDT selections made specifically by the EFDS; however, it would require changes to the RRP predictive model settings that would significantly increase the RRP FDR. The IRS can adjust the models to obtain the desired target fraud detection rate and the FDR; however, it must balance competing objectives of achieving high fraud detection without burdening taxpayers and IRS operations with excessive false detections. Since the RRP is not designed to detect the same returns as the EFDS, the only way the IRS could guarantee that the RRP would select the IDT returns selected by the EFDS would be to significantly lower its fraud risk score setting, resulting in significantly more legitimate taxpayers having their refunds delayed. With the EFDS IDT models contributing only 1.5 percent of the \$3.92 billion in total confirmed IDT selections not already detected by the RRP or the DDb in the 2015 Filing Season and the risk of increasing the RRP FDR (thus increasing taxpayer burden), we believe that the IRS made a reasonable decision to retire the EFDS IDT models.

In June 2015, the Executive Steering Committee approved a pilot of RRP Non-IDT filters. The IRS started the pilot in September 2015, *****2*****. The objectives of these filters was to detect fraud in which the perpetrator would be able to successfully pass the IDT filters. Since the Non-IDT filters were implemented so late in the 2015 Filing Season, the IRS made minimal changes to the Non-IDT filters for the 2016 Filing Season. The IRS ran the EFDS and the RRP Non-IDT filters in parallel for one full filing season. Figure 3 summarizes the number of Non-IDT returns selected by each individual fraud detection system as well as Non-IDT returns selected by both the RRP and the EFDS.

Figure 3: 2016 Filing Season Non-IDT Selections by System

Selection Source	Confirmed Non-IDT	Revenue Protected	FDR
RRP Selections	41,710	\$328 Million	55%
EFDS Selections	6,824	\$17 Million	70%
Selected by the RRP and the EFDS	7,249	\$55 Million	11%
Total ⁷	55,783	\$400 Million	

Source: IRS – Timeline of Return Review Program Expansion in Fraud Detection provided data.

⁷ The term total refers to all systemic selections. There are additional tax returns selected by manual processes that are not reflected in Figure 3.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

As Figure 3 shows, the RRP Non-IDT filters selected 41,710 fraudulent tax returns not selected by the EFDS, representing \$328 million in revenue protection. In comparison, the EFDS Non-IDT filters selected 6,824 fraudulent tax returns not selected by the RRP, representing \$17 million in revenue protected. Just as with IDT, we do not believe the relatively small number of Non-IDT tax returns selected by the EFDS warranted delaying the retirement of the EFDS Non-IDT filters after the 2016 Filing Season.

The Scope of the Return Review Program Was Expanded to Include Some Dependent Database Identity Theft Detection Functionality

Whereas the IRS designed the RRP as a replacement for the EFDS, the IRS views the RRP and the DDb as complimentary fraud detection systems with different functionality to combat IDT. As the IRS began to expand the RRP's capabilities to replace the EFDS, the IRS also looked for opportunities to leverage the RRP to identify more IDT at a lower FDR than the DDb. To accomplish this objective, the IRS performs a gap analysis at the conclusion of each filing season to identify gaps in IDT fraud detection between the DDb and the RRP. The gap analysis approach identifies the IDT tax returns selected by each of the DDb filters that were not selected by the RRP. Next, the IRS separates the confirmed DDb identity theft returns from the DDb false positive returns, allowing the IRS to focus on determining the confirmed DDb IDT tax returns it wants the RRP to capture in the subsequent filing season.

We reviewed the IRS's gap analysis from the 2015 to 2017 Filing Seasons. The June 2015 gap analysis showed that the DDb made 379,000 IDT selections not selected by the RRP, of which 138,978 were confirmed IDT. The IRS reviewed data for the 379,000 DDb selections, identifying the number of IDT returns that were a false positive, the number of returns that were confirmed IDT fraud, the FDR, and the percentage of the total DDb selections for each of the 27 DDb filters. From this analysis, the IRS developed three RRP strategies focusing first on the DDb filters with the highest volume of the fraud detection gap and the lowest FDR. The changes implemented in the RRP for the 2016 Filing Season resulted in a reduction of the original fraud detection gap to 70,474.⁸ However, the *****2***** in the DDb for the 2016 Filing Season that added 47,711 returns to the original confirmed IDT gap, resulting in a total of 118,185 (70,474 + 47,711) confirmed IDT tax returns.

The IRS performed a similar gap analysis of 2016 Filing Season results. In addition, the IRS made a few key decisions in June 2016: 1) the IRS decided that all new model development would be performed in the RRP; 2) the IRS would not retire the IDT filters in the DDb, resulting in the DDb filters continuing to make selections for the 2017 Filing Season; and 3) since the DDb was not being retired, RRP development resources would only be allocated to high-impact

⁸ Based on the IRS Gap Analysis dated June 27, 2016.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

filters.⁹ The IRS determined that only one DDb filter and one business rule met the high-impact criteria. Based on these results, the RRP collaborated with the Wage and Investment Division to deploy new RRP capabilities for the 2017 Filing Season that addressed the one DDb filter and one business rule that met the high-impact criteria.

We focused our review on the new *****2***** filter because the IRS analysis showed that this DDb filter accounted for the highest volume of DDb IDT selections in the 2016 Filing Season. For the 2016 Filing Season, the DDb filter selected 56,360 tax returns, of which 26,333 were confirmed as fraudulent (for an FDR of 68 percent).¹⁰

*****2*****
*****2*****. However, we reviewed a report from the 2017 Filing Season which demonstrated that the *****2***** are working and have accounted for 1,814 IDT selections totaling \$6 million in potential revenue protected as of March 17, 2017. This report indicates that the IRS is making further progress towards closing the IDT fraud detection gap between the DDb and the RRP. We agree with the June 2016 decisions because we believe they will stabilize the population of DDb-confirmed IDT selections not selected by the RRP, allowing the IRS to identify RRP enhancements to reduce the IDT fraud detection gap without simultaneously adding to it. In addition, by running the DDb and the RRP IDT in parallel, the IRS will increase its confidence that returns selected by both systems are in fact IDT.

The Systemic Utilization of Data Elements in the Return Review Program Was Expanded in the 2017 Filing Season

In March 2015, the IRS Commissioner convened a Security Summit meeting with officials from the IRS, private tax industries, and State departments of revenue to discuss common challenges and to work together to combat the emerging threat of IDT tax refund fraud. The summit worked to identify new steps to validate taxpayer and tax return information, increase information sharing between industry and Governments, and standardize sharing of suspected identity fraud information and analytics from the tax industry to identify fraud schemes and indicators of fraud patterns. The summit also established the creation of a working group that was responsible for identifying ways to help validate taxpayers. In subsequent meetings, the working group created 23 new Federal tax return data elements to be implemented by the IRS for the 2016 Filing Season. The new data elements were to help authenticate the taxpayer and detect

⁹ The IRS defines high impact as any filter meeting the following criteria: *****2*****
*****2*****
*****2***** It is also the same method applied to RRP filters.

¹⁰ The IRS Receiver Operating Characteristic Curve documentation shows an FDR of 64 percent. We attributed this discrepancy to the timing of the reports (the IRS Gap Analysis showing 68 percent was completed in June 2016; the Receiver Operating Characteristic Curve report was completed in August 2016). Testing of the filter estimated that the FDR would be 48 percent. However, the FDR was actually 68 percent.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

IDT tax refund fraud at the time of filing tax returns in support of revenue protection and fraud prevention strategies.

In a September 2016 report, TIGTA reported¹¹ that the IRS implemented the 23 new data elements into the RRP system based on the 2015 Security Summit, *Protecting Taxpayers from Identity Theft Tax Refund*. However, only three data elements were used systemically to filter returns and help identify tax refund fraud during the 2016 Filing Season. Further, the three data elements were implemented as business rules, as opposed to being implemented in the more robust RRP predictive models. For the remaining 20 data elements, there were insufficient historical data to create business rules that would enable systemic usage.

The Applications Development function informed us that they typically evaluate new data elements for ***2*** to determine their full potential or if they can be used systemically **2** ***2****. The evaluation also includes determining whether the data element will be consistent and stable over time. *****2*****. If the data elements the RRP uses are unstable or inconsistent *****2****, the IRS would risk implementing predictive models that falsely identified tax returns as fraudulent, thus increasing the RRP's FDR. *****2*****. The process the RRP uses to make the determination entails meeting with the Wage and Investment Division to develop an understanding of the data element and to identify any data quality problems. Fields that are deemed appropriate for modeling proceed to the next phases of the CRISP-DM, which are the iterative cycles of data preparation, modeling, and evaluation. During these phases, the IRS runs its proprietary RRP predictive model scoring technology over many weeks of tax return processing, ***2****. *****2*****. The proprietary scoring technology decides which data fields to use in the final RRP models.

The IRS significantly expanded the RRP's systemic use of new data elements in the 2017 Filing Season to detect IDT tax returns. In total, the RRP used ***2**** data elements in either a predictive model, a standalone filter, link return analysis,¹² or in a combination to expand IDT detection and decrease the FDR. We conducted our analysis of the **2** data elements to evaluate whether they met the criteria of consistent and stable over time. Our analysis concluded the following:

- The ***2*** data elements used met the IRS criteria for consistent and stable over time.
- Roughly half of the data elements are related to validating the taxpayer's identity. For example:

¹¹ TIGTA, Ref. No. 2016-20-062, *Filing Season 2016: Implementation of New Data Elements* (September 2016).

¹² Link Return Analysis is a tool that reveals patterns and relationships across masses of tax return data. It allows the RRP to identify clusters of returns that share characteristics indicative of tax fraud.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

- ***2*** data elements are related to the ***2***
 *****2*****
 *****2*****¹³ *****2*****
 *****2*****
 *****2*****.
- Other data elements to validate the taxpayer’s identity include ***2***
 *****2*****.
- The remainder of the data elements are related to ***2***
 *****2*****.
- *****2*****
 *****2*****
 *****2*****
 *****2*****.
- *****2*****
 *****2*****
 *****2*****
 *****2*****.

Having determined that ***2*** of the data elements met the IRS criteria for systemic use, we reviewed the IRS’s August 2016 Predictive Analytics Filing Season 2017 Status Report containing test results of the 2017 Filing Season RRP IDT predictive models to evaluate whether the data elements would expand IDT fraud detection and/or reduce the RRP FDR. The test results presented in Figure 4 indicate that the 2017 Filing Season RRP IDT predictive models improved the FDR over the 2016 Filing Season models at the same target detection rate. For example, we reviewed the RRP predictive model that targets electronic returns with more legitimate tax return data on the fraudulent return because the IRS indicated that this model will be the single largest driver of IDT detection in the 2017 Filing Season. This RRP model is specifically designed to distinguish between legitimate filers and identity thieves who have large amounts of data about the taxpayer to make the IDT tax return look legitimate.

¹³ These same ***2*** data elements apply to the primary taxpayer’s spouse.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Figure 4: RRP Predictive Model FDR Comparison

Target Detection Rate	2017 Filing Season FDR	2016 Filing Season FDR
2	7%	11%
2	12%	18%
2	25%	38%
2	50%	63%

Source: IRS Predictive Analytics Filing Season 2017 Status Report, dated August 2016.

As Figure 4 shows, at every listed target fraud detection rate, the 2017 Filing Season RRP model has a lower FDR compared to the equivalent 2016 Filing Season RRP model. A decrease in the FDR means fewer legitimate taxpayer returns are being falsely selected for fraud investigation.

While we typically do not test systems in the production environment, we did meet with the IRS in February 2017 to physically observe execution of the 2017 Filing Season RRP IDT predictive models in the production environment to verify systemic utilization of the data elements. Based on our observation, we confirmed that **2** data elements were used in the RRP predictive models, ***2*** were used in link return analysis/clusters, and ***2*** were used in standalone filters.¹⁴ In May 2017, the Application Development function provided us with data that illustrates the year-to-date impact of the ***2*** data elements used for the 2017 Filing Season.

Figure 5: Impact of the Authentication Data Elements for the 2017 Filing Season

Field	Impact
Link Return Analysis/Clustering	4,025 selections totaling \$18.8 million in tax return refund fraud stopped.
Filters to Detect IDT	275 tax returns selected totaling \$808,702 in tax return refunds stopped.
Filters to Reduce False Positives	119,160 tax return refunds excluded from IDT selection.

Source: Processing Year 2017 Impact of the Authentication Data Elements, dated May 2017.

*****2*****
 *****2*****. When a tax software company submits a return to the IRS, the tax software company creates *****2*****
 *****2*****
 *****2*****

¹⁴ The IRS used several of the data elements in more than one category.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

components to the RRP to ensure that cutover activities were completed. The IRS completed cutover activities in October 2016 as scheduled. The IRS stated that it reduced the annual operations cost of the EFDS by approximately \$650,000 as a result of retiring the EFDS Data Mining Models.

The two EFDS components that the IRS did not retire timely are the Affordable Care Act¹⁶ Fraud Models and Select and Characteristic Codes. Both components had a target retirement date of December 2016. At the time the IRS approved the EFDS Retirement Strategy, the IRS intended to add the Affordable Care Act Fraud Models into the EFDS. However, due to the volume of data and not wanting to degrade system performance, the IRS determined that it was best to store the information outside both the EFDS and the RRP. As an alternative to model development, the IRS created an interface between the Integrated Production Model and the EFDS. When we asked why the IRS did not create the interface between the Integrated Production Model and the RRP by December 2016, the IRS responded that it placed a higher priority on transitioning to Business Objects for RRP reporting. The IRS believed transitioning to Business Objects to improve reporting was of more importance because the IRS already had an existing interface between the Integrated Production Model and the EFDS. The IRS decided to move the target implementation date for the Integrated Production Model interface with the RRP to October 2017.

The Select and Characteristic Codes are used by the Wage and Investment business unit to determine which treatment stream tax returns flagged by the EFDS or the RRP are routed for further analysis. When we asked why the EFDS Select and Characteristic Codes component was not retired in December 2016, the IRS stated that the Select and Characteristic Codes implemented in the EFDS did not match the list planned for implementation in the RRP. As a result, the IRS wanted to do further analysis to ensure that the correct Select and Characteristic Codes are implemented in the RRP. The IRS is now scheduled to implement the Select and Characteristic Codes over two releases, with the first release scheduled for October 2017. The second release date has not been determined.

The remaining 11 of 19 components are related to the ECM and have retirement dates in December 2017 or December 2018. The IRS initiated the ECM as a separate project in January 2015, to be implemented in December 2018, thus enabling the IRS to retire the remaining EFDS components. However, in February 2017, the IRS suspended the ECM project. Since placing the ECM project on a suspension, the IRS assigned the responsibility of retiring the remaining EFDS components to a new entity, the EFDS Retirement Initiative. The EFDS Retirement Initiative will be overseen by the Information Technology organization's Enterprise Program Management Office. As of June 2017, the IRS stated that it has no indication of how long the ECM project suspension will last. The IRS suspended the ECM project due to

¹⁶ Pub. L. No. 111-148, 124 Stat. 119 (2010) (codified as amended in scattered sections of the Internal Revenue Code and 42 U.S.C.), as amended by the Health Care and Education Reconciliation Act of 2010, Pub. L. No. 111-152, 124 Stat. 1029.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

insufficient funding and staffing. In addition, the IRS determined that the software product selected for ECM cannot support an enterprise-wide deployment.

The IRS cannot completely retire the EFDS until all 19 components have been decommissioned. With the ECM project starting over with software selection, the IRS will likely miss the December 2018 target date to retire the remaining 11 EFDS components, which are dependent upon implementation of ECM. As a result, the IRS will continue to incur annual costs to operate and maintain the EFDS each filing season beyond the 2018 Filing Season. The IRS estimates that the annual operating and maintenance cost for the EFDS for the 2018 Filing Season will be \$13.9 million.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Appendix I

Detailed Objective, Scope, and Methodology

Our overall objective was to review the RRP¹ to determine if the system can identify all fraud currently identified by other existing fraud detection systems and to assess the EFDS retirement plans. To accomplish our objective, we:

- I. Determined the extent that EFDS and DDb fraud detection capabilities have been incorporated into the RRP.
 - A. Obtained a listing of the system capabilities, functionalities, and requirements for all three systems (EFDS, DDb, and RRP) for comparison to ensure that capabilities, functionalities, and requirements from the EFDS and DDb were incorporated in the RRP.
 - B. Interviewed the EFDS, DDb, and RRP subject matter experts to determine the process and procedures for incorporating capabilities, functionalities, and requirements into the RRP.
 - C. Determined the justifications and risks for any gaps and watch list items.
- II. Reviewed the IRS's plan for incorporating the 23 new Federal tax return data elements added for the 2016 Filing Season into the RRP predictive models for the 2017 Filing Season.
 - A. Obtained and reviewed documentation to determine which and how many of the 23 new Federal tax return data elements added for the 2016 Filing Season were incorporated into the RRP predictive models for the 2017 Filing Season.
 - B. Met with RRP subject matter experts to physically observe execution of the 2017 Filing Season RRP predictive models in the production environment to verify implementation of new data elements in the RRP predictive models.
- III. Evaluated the status of the EFDS retirement plans.
 - A. Obtained and reviewed the EFDS Retirement Strategy to determine the sequence and timing of events through eventual total EFDS shutdown.
 - B. Interviewed RRP and EFDS personnel and management to determine whether corrective actions to a prior TIGTA recommendation with regards to documenting a retirement plan were completed.

¹ See Appendix IV for a glossary of terms.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

- C. Determined the plan the IRS had in place to ensure that the overall performance in identifying refund fraud was not compromised with the retiring of the EFDS system.
- D. Reviewed and evaluated shutdown procedures used for those elements of the EFDS and the DDb that had transitioned to the RRP prior to the 2017 Filing Season (*i.e.*, IDT models and Non-IDT models).

Internal controls methodology

Internal controls relate to management's plans, methods, and procedures used to meet their mission, goals, and objectives. Internal controls include the processes and procedures for planning, organizing, directing, and controlling program operations. We determined that the following internal controls were relevant to our audit objective: the Internal Revenue Manual and related IRS guidelines and the processes followed in the development and retirement of information technology projects. We evaluated these controls by conducting interviews with management and staff and reviewing documentation. Documents reviewed include the *EFDS to RRP Traceability*; the EFDS Retirement Strategy and associated deliverables; and other documents that provided evidence of whether the IRS is adequately managing RRP systems development risks during the ongoing retirement of the EFDS.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Appendix II

Major Contributors to This Report

Danny R. Verneuille, Assistant Inspector General for Audit (Security and Information
Technology Services)
John L. Ledford, Director
Michael Mohrman, Audit Manager
Mark Carder, Lead Auditor
Chinita Coates, Auditor



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Appendix III

Report Distribution List

Commissioner
Office of the Commissioner – Attn: Chief of Staff
Deputy Commissioner for Operations Support
Deputy Chief Information Officer for Operations
Associate Chief Information Officer, Enterprise Information Technology Program Management
Office
Director, Office of Audit Coordination



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Appendix IV

Glossary of Terms

Term	Definition
Data Breach	The potential or actual loss of control, compromise, unauthorized disclosure, unauthorized acquisition, unauthorized access, or any similar term referring to situations in which persons other than authorized users and for other than authorized purposes have access or potential access to Personally Identifiable Information, whether physical or electronic.
Data Element	The smallest named item of data that conveys meaningful information or condenses a lengthy description into a short code.
Dependent Database	A risk-based audit selection tool used by the IRS to identify tax returns for audit.
Electronic Fraud Detection System	An automated system used to maximize fraud detection at the time tax returns are filed to eliminate the issuance of questionable refunds.
Filing Season	The period from January through mid-April when most individual income tax returns are filed.
Filters	A mechanism in the IRS fraud detection systems that selects potential fraudulent returns and routes them for appropriate IRS analysis/treatment.
Identity Theft	For the purpose of tax fraud, occurs when an individual uses another person's name and Taxpayer Identification Number (generally a Social Security Number) to file a fraudulent tax return to obtain a fraudulent tax refund.
Integrated Production Model	A data warehouse that consolidates information from a variety of internal, and some external, sources. That data are made available to a variety of downstream security-certified systems for use in conducting analysis, case selection, and report preparation.



*The Return Review Program Increases Fraud Detection;
However, Full Retirement of the Electronic Fraud
Detection System Will Be Delayed*

Term	Definition
Predictive Models	Use known results to develop (or train) a model that can be used to predict values for different or new data. The modeling results in predictions that represent a probability of the target variable (e.g., revenue) based on estimated significance from a set of input variables.
Processing Year	The calendar year in which the tax return or document is processed by the IRS.
Requirements	Describes a condition or capability to which a system must conform, either derived directly from user needs or stated in a contract, standard, specification, or other formally imposed document. A desired feature, property, or behavior of a system.
Return Review Program	A system the IRS uses to identify potentially fraudulent electronically filed tax returns. It enhances the IRS's capabilities to detect, resolve, and prevent criminal and civil noncompliance and reduces issuance of fraudulent tax refunds.
Traceability (also known as Crosswalk Traceability)	The creation and maintenance of a discernable association among two logical entities such as requirements, system elements, verifications, or tasks.