TREASURY INSPECTOR GENERAL FOR TAX ADMINISTRATION

The Integrated Production Model
Increases Data Access Efficiency; However,
Access Controls and
Data Validation Could Be Improved

July 29, 2016

Reference Number: 2016-20-058

THE INTEGRATED PRODUCTION MODEL INCREASES DATA ACCESS EFFICIENCY; HOWEVER, ACCESS CONTROLS AND DATA VALIDATION COULD BE IMPROVED

Highlights

Final Report issued on July 29, 2016

Highlights of Reference Number: 2016-20-058 to the Internal Revenue Service Chief Information Officer.

IMPACT ON TAXPAYERS

The Integrated Production Model (IPM) replaced two similar systems in February 2007. The IPM now provides a single point of access to current and historical core taxpayer data in a centralized repository. The consolidation of systems reduces the need for redundant databases. The accuracy, completeness, and reliability of data on the IPM are essential to the IRS and its tax administration mission.

WHY TIGTA DID THE AUDIT

This audit is included in our Fiscal Year 2016 Annual Audit Plan and addresses the major management challenge of Improving Tax Systems and Online Services. The overall objective of this review was to determine whether the data within the IPM are accurate, complete, and reliable.

WHAT TIGTA FOUND

The IPM system is meeting IRS business needs and has improved the efficiency of data access via a singular data repository that has taken over the processing load of two separate database systems. However, access controls were not documented, and TIGTA was unable to definitively verify that the IPM pulls data from only designated source systems. The overall design of the IPM system allows access controls to be managed by a system external to the IPM database—the Big Data Analytics program. In addition, 14 (77 percent) of 18 IPM source systems reviewed have no validation of data for accuracy, completeness, and reliability. The

IPM database acts as a data repository, and there are no controls to validate received data. The IPM assumes that all data received are accurate, complete, and reliable.

WHAT TIGTA RECOMMENDED

TIGTA recommended that the Chief Technology Officer: 1) establish updated access authorization documentation for the IPM in each of the source systems; 2) conduct a periodic review of access control lists to verify that all systems are current, authorized, and documented; and 3) implement data validation steps for information pulled into the IPM to ensure that it is accurate, complete, and reliable.

The IRS agreed with two recommendations and disagreed with one recommendation. The IRS agreed to update IPM access authorization documentation and to conduct annual validation of IPM service accounts as part of its filing season update process. The IRS disagreed that additional data validation steps should be implemented for the IPM, stating that appropriate controls are currently in place.

TIGTA maintains the controls currently in place are not sufficient to ensure data reliability in the IPM. Without adequate data validation for the IRS's core taxpayer data store, undetected errors could result in corrupt data being provided to other applications that support case identification, selection, prioritization, delivery, and reporting.

DEPARTMENT OF THE TREASURY

WASHINGTON, D.C. 20220

TREASURY INSPECTOR GENERAL
FOR TAX ADMINISTRATION

July 29, 2016

MEMORANDUM FOR CHIEF INFORMATION OFFICER

FROM: Michael E. McKenney
Deputy Inspector General for Audit

SUBJECT: Final Audit Report – The Integrated Production Model Increases Data Access Efficiency; However, Access Controls and Data Validation Could Be Improved (Audit # 201520019)

This report presents the results of our review of the Integrated Production Model. The overall objective of this review was to determine whether the data within the IPM are accurate, complete, and reliable. This review is included in the Treasury Inspector General for Tax Administration's Fiscal Year 2016 Annual Audit Plan and addresses the major management challenge of Improving Tax Systems and Online Services.

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

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

Table of Contents

Abbreviations

IPM	Integrated Production Model
IRS	Internal Revenue Service
NIST	National Institute of Standards and Technology

Background

The Internal Revenue Service (IRS) implemented the Integrated Production Model (IPM) in February 2007. The IPM is a centralized analytical data store that provides a single point of access to core taxpayer data (such as taxpayer accounts and tax returns) and other specific data used by a wide range of IRS business applications to support case identification, selection, prioritization, delivery, and reporting.

The IPM system was designed as a replacement for two legacy systems—the Enterprise Data Warehouse Business Filers Model and the Enterprise Data Warehouse Individual Filers Model. These legacy systems previously processed data over the course of several hours per query and required users to run queries in both systems separately in order to collect the data required to meet business needs. The IPM system was designed to benefit the IRS by providing multiple IRS business organizations access to current and historical taxpayer data. Once implemented, the IPM database[1] eliminated the need for separate redundant databases as well as created a single point of contact for IRS data.

The IPM system is a database application and subcomponent segment of the IRS Big Data Analytics program. Big Data Analytics is a General Support System that provides an enterprise-level analytical platform using the Greenplum data warehouse application. Big Data Analytics is considered an appliance that will provide the IRS with the ability to conduct advanced analytics and low-latency data processing as well as in-depth analysis of data. This system will be able to handle datasets and process analytics in a fast environment. Further, Big Data Analytics will serve to perform advanced data analysis that can facilitate IRS audit selections, analyzing taxpayer filings, and more.

The technology enabling these functions is the Massively Parallel Processing architecture that has been designed for business intelligence and analytical processing. In this architecture, data is automatically partitioned across multiple segment servers, and each segment owns and manages a distinct portion of the overall data. Big Data Analytics manages the user interface and other administrative elements of data access for its segment servers, including the IPM system. The IPM uses workflows programmed in third-party software (Informatica) to pull data from source systems. See Figure 1 for an illustration of the overall IPM system design.

---

[1] See Appendix VI for a glossary of terms.

Figure 1: Integrated Production Model – Design Overview
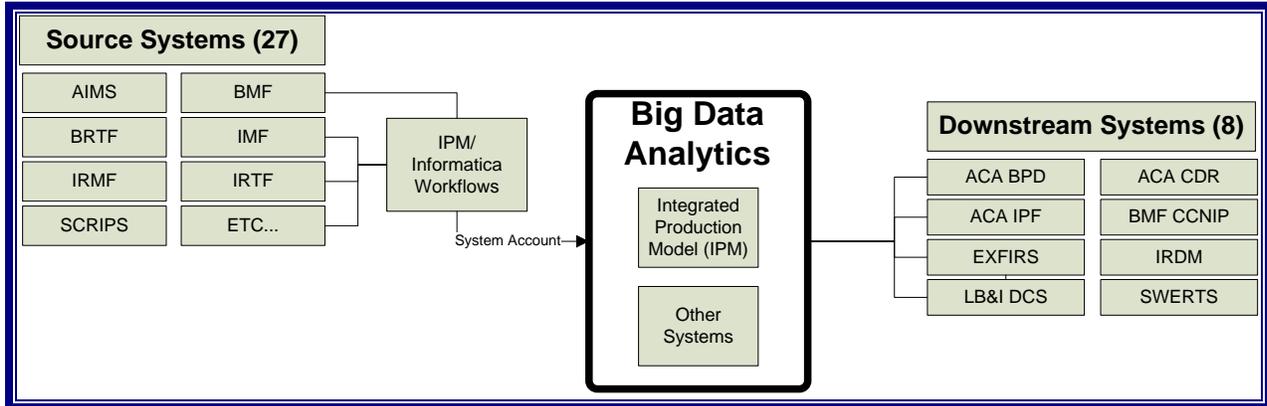
Source: Our analysis of the IPM Design Specification Report and discussions with the IPM program management team. See Appendices IV and V for a list of system acronyms used in this figure.

This review was performed at the IPM project office located in the New Carrollton Federal Building in Lanham, Maryland, during the period September 2015 through February 2016. 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.

Results of Review

The Integrated Production Model Is Meeting the Business Needs of Its Downstream Users

As noted in the President's proposed budget for Fiscal Year 2017, duplicative functions and services have made Government less effective and wasted taxpayer dollars. Improving efficiency in Federal programs is a core component of the President's Management Agenda. Among other project management principles, organizations should consider the effect of project decisions on the subsequent recurring cost of using, maintaining, and supporting the product or service.[2]

The IPM is an improvement over its predecessors—the Enterprise Data Warehouse Business Filers Model and the Enterprise Data Warehouse Individual Filers Model—in terms of speed and in serving the business needs of its users. Both predecessor systems, retired in 2007, extracted data from various source systems and then reformatted the data into relational databases. These databases were then used to identify and measure tax noncompliance from business or individual filers, respectively. The IPM now provides a single point of access to current and historical core taxpayer data and its centralized repository. The consolidated system reduced the need for multiple queries by aggregating the data and making corporate and individual taxpayer data available in a single location. We interviewed subject matter experts and business owners from seven of eight downstream systems, who told us that they are satisfied with the services and data provided from the IPM database. Personnel from the Affordable Care Act Program Management Office, Small Business/Self-Employed Division, and Large Business and International Division indicated that the IPM data meet their business needs. The downstream system owners also stated that there have been no significant issues with the accuracy, completeness, or reliability of data received from IPM.

We reviewed all closed tickets, *i.e.*, 268 incident tickets and two problem tickets, relating to the IPM system between September 1, 2014, and September 1, 2015, and determined that there were no significant instances of failures with the IPM database.

The IRS Did Not Follow Established Processes for Authorizing and Documenting Access Controls

The National Institute of Standards and Technology (NIST) defines access control lists as a register of users (including groups, machines, processes) who have been given permission to use

---

[2] Project Management Institute, *A Guide to the Project Management Body of Knowledge* (5th Edition, 2013).

a particular system resource and the types of access they have been permitted.[3] The NIST also provides guidance regarding the principle of least privilege, which means allowing only authorized accesses for users that are necessary to accomplish assigned tasks in accordance with organizational missions and business functions.[4] Organizations employ least privilege for specific duties and information systems. Organizations consider the creation of additional processes, roles, and information system accounts, as necessary, to achieve least privilege. Further, the need for certain assigned user privileges may change over time, reflecting changes in organizational missions, business function, operation environment, technologies, or threat. The IPM project team told us that there is a formalized process in place for gaining access to the IPM that is controlled by owners of the Big Data Analytics program upon which the IPM resides. The process consists of completing documentation that outlines the business need for access as well as the specific level of access that is required. The Big Data Analytics system owners are responsible for maintaining compliance with documented access requirements.

The IPM Design Specification Report, released June 1, 2015, indicates that there are 27 source systems.[5] To test whether the information provided in the Design Specification Report is valid and accurate, we requested a system-generated or automated list of IPM source systems. According to Greenplum technical guidance, the Greenplum data warehouse application is capable of identifying authorized systems. The IPM project team, in conjunction with the Informatica/Greenplum team, could not provide documentation showing that the IPM accesses only these 27 source systems. Greenplum data warehouse application staff provided logs intended to illustrate the transmission of this information, but our review of the logs showed no unique identifiers that can be used to identify the destination of the actions taken in the logs. In an effort to manually verify that workflows are pulled only from the designated 27 source systems, we also reviewed three additional sets of logs from the Informatica interconnection with the IPM. Our detailed review of the logs did not yield any specific identifiers that allowed for a comprehensive list of source systems to be documented.

Data is communicated between the source systems and the IPM database via 27 read-only service accounts (one account per source system). The IPM is authorized within each of the source system's boundary by a read-only service account that pulls data from that specific system. Thus, the IPM is essentially a user of each of the 27 source systems. Users must request access to a system via the IRS Online 5081 application, where relevant system owners then authorize and approve those requests. We attempted to verify this process and confirm that access controls are in place and procedures were followed during IPM system development. We requested copies of authorization request and approval documents for the IPM service accounts on each of the 27 source systems. After several weeks had elapsed, IPM system administrators informed us that access request documentation was deleted in Calendar Year 2010. Therefore,

---

[3] NIST, Special Publication 800-12, *An Introduction to Computer Security: The NIST Handbook* (October 1995).
[4] NIST, Special Publication 800-53, *Security and Privacy Controls for Federal Information Systems and Organizations* (April 2013).
[5] See Appendix IV for a complete list of IPM source systems.

we were provided no documentation or confirmation that the IPM is authorized to access the designated 27 source systems. Combined with our work analyzing audit logs in which we could not distinguish unique workflows from specific systems, we could also not verify that the IPM pulls data only from these systems alone.

Without documentation and adequate management of access controls and without the ability to review system audit logs to verify unique system access, the IRS cannot be sure which systems the IPM pulls data from. Furthermore, according to the NIST, periodic review of assigned user privileges is necessary to determine if the rationale for assigning such privileges remains valid. In this instance, at least 27 systems showed evidence of incomplete, inaccurate, and outdated documentation for user access. As systems such as Big Data Analytics and its various components grow and additional functionality is added, maintaining access controls and following established processes becomes more critical as systems change and become more complex.

Recommendations

Recommendation 1: The Chief Technology Officer should establish updated access authorization documentation for the IPM in each of the source systems.

> **Management's Response**: The IRS agreed with this recommendation. The IPM project team will resubmit records for each source system service account to document the access authorization.

Recommendation 2: The Chief Technology Officer should conduct a periodic review of access control lists to verify that all systems are current, authorized, and documented.

> **Management's Response**: The IRS agreed with this recommendation. The IPM project team will establish an annual validation of IPM service accounts as part of its filing season update process and will document service account information in Interface Control Documents for source systems.

The Integrated Production Model Database Utilizes Data That Have Not Been Validated for Accuracy, Completeness, or Reliability

The Government Accountability Office defines data reliability as data that are reasonably complete and accurate, meet the intended needs to the user, and are not subject to inappropriate alteration.[6] Accuracy is characterized in greater detail by the extent that recorded data reflects the actual underlying information. Completeness is defined further by the extent that records are present and the fields in each record are properly populated.

---

[6] Government Accountability Office, GAO-09-680G, *Assessing the Reliability of Computer-Processed Data* (July 2009).

The IRS considers the IPM database to be a data store and considers the data valid as received; therefore, there is no obligation to verify the data. As such, the IPM system does not perform any validation of received data other than a record count to ensure that the number of files expected was received. According to industry standards for information technology management and governance, controls over information architecture help provide reliable and consistent information to integrate applications into business processes.[7] One such control is measured by the frequency of data validation activities to ensure the integrity and consistency of all data stored in electronic form, such as databases, data warehouses, and data archives. As shown in Figure 2, we selected 18 of 27 source systems for review via a judgmental sample[8] based on the system's apparent impact on tax administration processes. We interviewed the system owners, and 10 (55 percent) of 18 told us that they consider the source to be trusted and do not conduct any data validation. An additional three (17 percent) of the 18 systems use data from a publicly available site and, as a result, do not conduct any data validation. These sources combined result in 13 (72 percent) of 18 source systems conducting no data validation and assuming that data are valid upon receipt from the source. Three (17 percent) of the 18 systems conduct only minimal validation to ensure that field data types are received as expected (such as alpha or numeric characters). Of the two remaining systems, one conducts validation of only the Taxpayer Identification Number, and the other completes no validation of any data – and its data do not come from a trusted source system.

Figure 2: IPM Source System – Data Validation Assessment
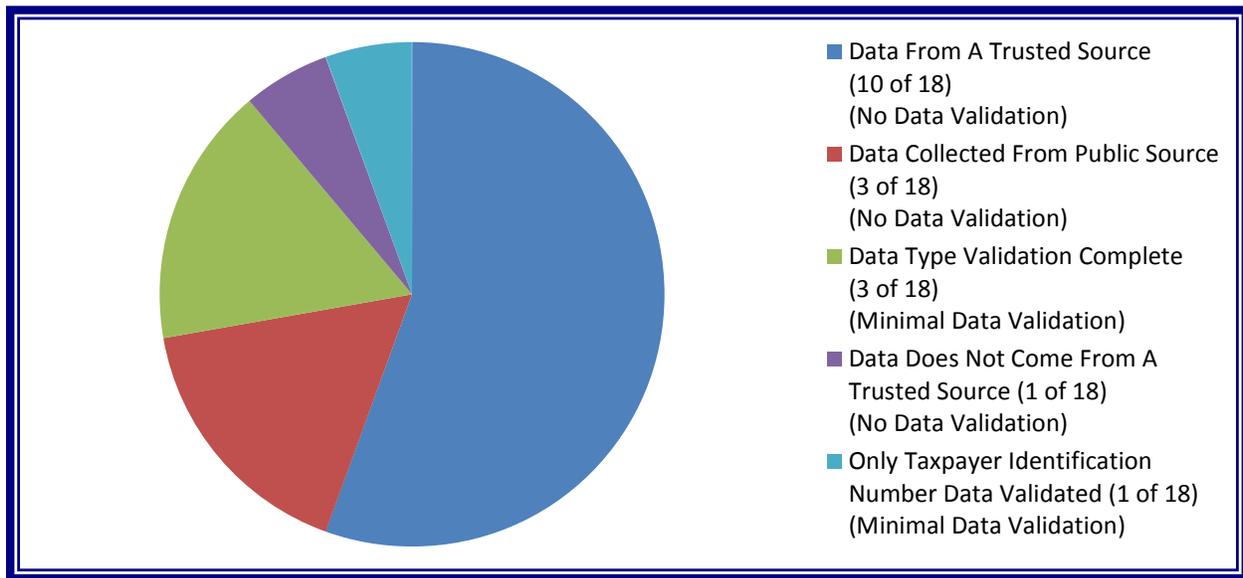
Source: Our analysis of data validation processes for 18 of 27 IPM source systems.

---

[7] Information Systems Audit and Control Association, *Control Objectives for Information Technology* (COBIT®5), 2012.

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

The IPM database functions as only a data repository; as a result, the IPM project team places the responsibility of ensuring the accuracy, completeness, and reliability of the data received on the upstream systems. According to the IPM Business System Report and the IPM project team, the IPM project does not have any requirements to validate data in the production environment.

Certain types of storage errors can go completely unreported and undetected. These undetected errors can result in corrupt data being provided to other applications. When applications and other systems receive this corrupted data, it could lead to inaccurate analytical results. Without data validation, the risk increases that inaccurate, incomplete, or unreliable data could be passed from source systems through the IPM database to downstream systems that use the data for case selection. Using inaccurate, incomplete, or unreliable data could result in the erroneous selection of cases for audit or further review. Erroneous case selection could result in the misallocation of limited IRS personnel and the IRS missing larger tax violations for which collections are necessary.

Recommendation

Recommendation 3: The Chief Technology Officer should implement additional data validation steps for data pulled into the IPM to ensure that information is accurate, complete, and reliable.

> **Management's Response:** The IRS disagreed with this recommendation. The IRS stated that it agreed with the facts related to data validation of the IPM data and its source systems, but it believes that appropriate controls are currently in place to ensure the accuracy, completeness, and reliability of these systems.

> **Office of Audit Comment:** Based on the results of our audit, we maintain that the controls currently in place are not sufficient to ensure data reliability in the IPM. Without adequate data validation for the IRS's core taxpayer data store, undetected errors can result in corrupt data being provided to other applications that support case identification, selection, prioritization, delivery, and reporting.

Detailed Objective, Scope, and Methodology

Our overall objective was to determine whether the data within the IPM are accurate,[1] complete, and reliable. To accomplish our objective, we:

I. Determined whether the IRS performed data reliability assessments on source systems that provide data to the IPM.

   A. We identified a list of all systems that provide data to the IPM.

   B. We obtained and reviewed criteria and procedures for ensuring data reliability on these source systems.

   C. We interviewed system owners to determine whether data reliability assessments were performed on these source systems. We selected 18 of 27 systems to review via a judgmental sample[2] based on the impact to IRS tax administration processes.

II. Evaluated the IRS process for validating the accuracy, completeness, and reliability of the data in the IPM.

   A. We obtained and reviewed criteria and procedures for validating the accuracy, completeness, and reliability of the data in the IPM.

   B. We interviewed IRS officials to discuss the process used to validate the accuracy, completeness, and reliability of the data in the IPM.

III. Determined whether there are any data reliability issues in the IPM.

   A. We obtained and analyzed an extract of problem tickets related to the IPM in order to identify any existing data issues within the IPM.

   B. We interviewed relevant IRS personnel about the results of our ticket report analysis.

   C. We obtained and analyzed supporting documentation and resolutions for each of the IPM problem tickets identified.

Internal controls methodology

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

---

[1] See Appendix VI for a glossary of terms.
[2] 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: the NIST guidance which requires the IRS to provide assurance that the accuracy, completeness, and reliability of data are maintained. Controls also indicate that data validation steps are to be taken in order to control the integrity of data residing in data systems. We evaluated these controls by reviewing IRS documentation and conducting interviews with IRS subject matter experts in order to understand the overall design as well as internal controls in place related to the accuracy, completeness, and reliability of data that reside within the IPM database. We also reviewed system-generated reports from the IPM as well as problem tickets related to the IPM.

Major Contributors to This Report

Danny R. Verneuille, Acting Assistant Inspector General for Audit (Security and Information Technology Services)
Myron L. Gulley, Acting Director
Jena R. Whitley, Audit Manager
Nicholas Reyes, Lead Auditor
Adaline B. LaForest, Senior Auditor
Michael T. Mohrman, Senior Auditor

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, Applications Development
Associate Chief Information Officer, Strategy and Planning
Director, Office of Audit Coordination

The Integrated Production Model Increases
Data Access Efficiency; However, Access Controls
and Data Validation Could Be Improved

Appendix IV

Integrated Production Model Source Systems

System Number	System Acronym	System Name	System Owners Interviewed (Y or N)
1	ACV	Affordable Care Act Compliance Validation	N
2	AIMS	Audit Information Management System	Y
3	AIR	ACA Information Returns (Processing)	N
4	ALEF	Allowable Living Expenses File	N
5	BMF	Business Master File	Y
6	BRTF	Business Return Transaction File	Y
7	CAWR	Combined Annual Wage Reporting	Y
8	CB	Census Bureau	Y
9	CDE	Compliance Data Environment	Y
10	CDR	Coverage Data Repository	N
11	CP2100BWH	Backup Withholding	Y
12	CTW	Chapter Three Withholding	Y
13	DUPTIN	Duplicate Taxpayer Identification Number	Y
14	EPMF	Employee Plans Master File	Y
15	ETEP	Employment Tax Examination Program	N
16	IMF	Individual Master File	Y
17	IRDB	Information Returns Database	N
18	IRMF	Information Returns Master File	Y
19	IRTF	Individual Return Transaction File	Y
20	NAICS	North American Industry Classification System	N
21	PMF	Payer Master File	Y

System Number	System Acronym	System Name	System Owners Interviewed (Y or N)
22	REIC	Research Earned Income Credit	N
23	SCRIPS	Service Center Recognition Image Processing System	Y
24	SS8ICP	Social Security Integrated Case Processing	Y
25	SSA-DM1	Social Security Administration	N
26	ZIP	Zone Improvement Plan	Y
27	ZIP_POD	Zone Improvement Plan Post of Duty	Y

**The Integrated Production Model Increases
Data Access Efficiency; However, Access Controls
and Data Validation Could Be Improved**

Appendix V

Integrated Production Model Downstream Systems

System Number	System Acronym	System Name	System Owners Interviewed (Y or N)
1	ACA BPD	Affordable Care Act Branded Prescription Drug	Y
2	ACA CDR	Affordable Care Act Coverage Data Repository	N
3	ACA IPF	Affordable Care Act Insurance Provider Fee	Y
4	BMF CCNIP	Business Master File Case Creation Nonfiler Identification Process	Y
5	EXFIRS	Excise Files Information Retrieval System	Y
6	IRDM	Information Reporting and Document Matching – Exam Case Selection	Y
7	LB&I DCS	Large Business and International Data Capture System	Y
8	SWERTS	Service-Wide Employment Tax Research System – Specialty Examination Policy and Quality	Y

**The Integrated Production Model Increases
Data Access Efficiency; However, Access Controls
and Data Validation Could Be Improved**

Appendix VI

Glossary of Terms

Term	Definition
Accuracy (of data)	The extent that recorded data reflect the actual underlying information.
Appliance	A computing device that provides predefined services and has its underlying operating software hidden beneath an application-specific interface.
Application	A software program hosted by an information system.
Authorization	Access privileges granted to a user, program, or process or the act of granting those privileges.
Completeness (of data)	The extent that records are present and the fields in each record are properly populated.
Database	A computer system with a means of storing information in such a way that information can be retrieved.
Design Specification Report	Documents the logical and physical design of a proposed solution from all applicable perspectives.
Downstream	Any system or systems that are taking data from a data source.
Fiscal Year	Any yearly accounting period, regardless of its relationship to a calendar year. The Federal Government's fiscal year begins on October 1 and ends on September 30.
Incident Ticket	Any event that is not part of the standard operation of a service and that causes, or may cause, an interruption to or a reduction in the quality of that service. The stated Information Technology Infrastructure Library objective is to restore normal operations as quickly as possible with the least possible impact on either the business or the user at a cost-effective price.
Informatica	An application system that functions as an intermediary between computer systems. This system can be responsible for the Extract Transform and Load process.

Term	Definition
Information System	A discrete set of information resources organized for the collection, processing, maintenance, use, sharing, dissemination, or disposition of information.
Least Privilege	The security objective of granting users only those accesses they need to perform their official duties.
Low Latency	A data solution with a potentially near-real-time query response.
National Institute of Standards and Technology	A part of the Department of Commerce that is tasked with creating standards for technology.
Online 5081	A web-based application that allows users to request access, modify existing accounts, reset passwords, and request deletion of accounts when access is no longer needed to specific systems. The application also allows the IRS to track user access history, generate reports, and document an audit trail of user actions
Problem Ticket	The unknown cause of one or more incidents often identified as a result of multiple similar incidents.
Relational Database	A database that presents information in tables with rows and columns.
Reliability (of data)	Data that are reasonably complete and accurate, meet the intended needs of the user, and are not subject to inappropriate alteration.
Upstream	Any system or systems that are considered a data source and are passing their data into another system.

Management's Response to the Draft Report

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

CHIEF TECHNOLOGY OFFICER

JUN 2 7 2016

MEMORANDUM FOR DEPUTY INSPECTOR GENERAL FOR AUDIT

FROM: Terence V. Milholland
           Chief Technology Officer

SUBJECT: The Integrated Production Model Increases Data
           Access Efficiency; However, Access Controls and
           Data Validation Could Be Improved (Audit
           #201520019) (e-trak # 2016-82545)

Thank you for the opportunity to review your draft audit report. The Internal Revenue
Service is committed to improving our overall efficiency. This is especially true with the
development of new systems, such as the Integrated Production Model (IPM), which is
the subject of this audit.

Thank you for recognizing our successful development and implementation of the IPM.
The system continues to improve the efficiency of the Services' data access and is a
cornerstone of our move to a data-centric organization.

We fully agree with the recommendations related to access controls and have provided
corrective actions to address TIGTA's findings. While we acknowledge the references in
your report related to the validation of IPM data, it is our contention that our Systems of
Record, as defined by The Privacy Act of 1974 and IRM guidelines, are accurate,
complete and reliable in serving as the basis for all refund, notice and compliance
actions.

The bulk of data loaded into IPM is tax account, tax return and information return data.
IPM pulls this data from the authoritative Systems of Record such as Individual
Masterfile (IMF), Business Masterfile (BMF) and Information Return Masterfile (IRMF).
While specific processes known as "data validation" may not be in place for these
systems, they contain extensive balance and control processes, such as run-to-run
balancing, to ensure they are complete. Systems containing financial data also contain
extensive financial balancing and reporting (which are subject to GAO audit) to ensure
that they are accurate.

2

Internal data checks in these systems include valid data ranges, restricted values, data type and business rules for acceptable values. Failures in these checks can lead to unpostables or system halts in which an incoming record is not permitted to post until the bad record is either removed or corrected.

We also believe that current IPM data validation processes are appropriate. As discussed with the audit team, IPM staff perform manual validation after each successful load to ensure that row counts in IPM match source system counts – verifying that IPM is complete. Since the audit was conducted, this process has been automated and will immediately notify staff if an imbalance is detected.

During annual IPM System Acceptance Testing, Enterprise Systems Testing staff conduct field-by-field data comparisons between IPM and source systems to ensure the accuracy of IPM data. All data discrepancy defects are assessed and appropriately dispositioned before updated IPM workflows are promoted to production.

As noted in the report, the audit team interviewed owners of downstream systems that use IPM as a data source, and none reported significant issues with the accuracy, completeness or reliability of IPM data.

The IRS values the analysis and recommendations your organization provides to improve our IT systems and business processes. If you have any questions, please contact me at (240) 613-9373, or contact Karen Mayr on (202) 368-8396.

Attachment

The Integrated Production Model Increases Data Access Efficiency; However, Access Control and Data Validation Could Be Improved (201520019)

RECOMMENDATION #1: The Chief Technology Officer should establish updated access authorization documentation for IPM in each of the source systems.

CORRECTIVE ACTION #1: We agree with the recommendation. IPM will resubmit OL5081 records for each source system service account to document the access authorization.

IMPLEMENTATION DATE: October 15, 2016

RESPONSIBLE OFFICIALS: Associate Chief Information Officer, Applications Development; Director, Data Delivery Services

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 Technology Officer should conduct a periodic review of access control lists to verify all systems are current, authorized, and documented

CORRECTIVE ACTION #2: We agree with the recommendation. IPM will establish an annual validation of IPM service accounts as part of its filing season update process and will document service account information in Interface Control Documents for source systems.

IMPLEMENTATION DATE: February 15, 2017

RESPONSIBLE OFFICIALS: Associate Chief Information Officer, Applications Development; Director, Data Delivery Services

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.

1

Attachment

*The Integrated Production Model Increases Data Access Efficiency; However, Access
Control and Data Validation Could Be Improved (201520019)*

RECOMMENDATION #3: The Chief Technology Officer should implement additional data validation steps for data pulled into IPM to ensure that information is accurate, complete, and reliable.

CORRECTIVE ACTION #3: We disagree with the recommendation. While we agree with the facts related to data validation of IPM and its source systems, we believe that appropriate controls are currently in place to ensure the accuracy, completeness and reliability of these systems.

IMPLEMENTATION DATE: N/A

RESPONSIBLE OFFICIALS: Associate Chief Information Officer, Applications Development; Director, Data Delivery Services

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.

2