

Date of Approval: 11/03/2025
Questionnaire Number: 2659

Basic Information/Executive Summary

What is the name of your project (system, database, pilot, product, survey, social media site, etc.)?

Information Returns Master File Modernization

Acronym:

IRMF Mod

Business Unit

Information Technology

Preparer

For Official Use Only

Subject Matter Expert

For Official Use Only

Program Manager

For Official Use Only

Designated Executive Representative

For Official Use Only

Executive Sponsor

For Official Use Only

Executive Summary: Provide a clear and concise description of your project and how it will allow the IRS to achieve its mission.

The mission of Information Returns Master File Modernization (IRMF Mod) is to modernize the highly complex legacy IRMF Master File (IRMF) ecosystem responsible for storing and processing information returns submitted to IRS including the following information: income, payments, retirements contributions withholding etc. IRMF is primarily used to match third-party reports with taxpayer filings to detect underreporting, non-filing, or fraud. IRMF extracts are used by both internal and external users; Internal users include Large Business and International (LB&I), Privacy, Governmental Liaison, and Disclosure (PGL&D), Small Business / Self-Employed (SB/SE), Tax-Exempt Government Entities (TEGE), National Headquarters (NHQ), Research, Applied Analytics and Statistics (RAAS), and Information Technology (IT). External users include Census and dozens of various State and Local Government agencies via Privacy, Governmental Liaison, and Disclosure (PGL&D). The modernization of the IRS

Information Returns Master File (IRMF) is a critical initiative aimed at transforming a legacy, batch-based system into a scalable, real-time, cloud-native platform. This effort supports key business objectives such as improving taxpayer compliance through faster and more accurate matching of third-party information returns, enabling rapid adaptation to changing tax laws, and enhancing fraud detection capabilities. By integrating IRMF data with modern analytics platforms like Databricks, the IRS can empower analysts with timely insights while reducing operational overhead and system risk. The modernization also aligns with strategic goals by replacing outdated Assembler Language Code (ALC) / Common Business Oriented Language (COBOL) systems with event-driven microservices and enabling secure, role-based data access. Ultimately, this transformation will enhance the taxpayer experience, increase voluntary compliance, and strengthen the IRS's ability to protect revenue and enforce tax law effectively. In summary, the IRMF Modernization program will:

- Introduce a cloud-based, vendor-built solution that will implement/enable Event-Driven Architecture (EDA) for near real-time processing using the information from Enterprise Data Platform (EDP) be the authoritative data source for business tax information,
- Facilitate retirement of legacy processes/systems; and
- Empower the Agency to more quickly deliver modern features and respond to requested changes.

Personally Identifiable Information (PII)

Will this project use, collect, receive, display, store, maintain, or disseminate any type of Sensitive but Unclassified (SBU), Personally Identifiable Information (PII), or Federal Tax Information (FTI)?

Yes

Please explain in detail how this project uses sensitive data from inception to destruction (data lifecycle).

The IRP system receives data submitted by employers and other third parties (payers) reporting taxpayer income such as wages, pensions, interest and dividends paid during the tax year. This information is validated and stored in the Information Return Master File (IRP: IRMF). IRP provides information technology support to complementary compliance functions allowing the Internal Revenue Service (IRS) to effectively administer the U. S. tax system. Data is sent to the IRP system and its subordinate applications via a variety of pipeline processes. IRP then attempts to validate all data received, but regardless of the outcome, all data received will be stored on one of several data stores that IRP controls. Upon receipt of an authorized request, IRP performs a data extract and sends requested information internally and externally via GSS-21.

Please select all types of Sensitive but Unclassified data (SBU)/Personally Identifiable Information (PII)/Federal Tax Information (FTI) that this project uses.

Address
Employer Identification Number
Employment Information
Federal Tax Information (FTI)
Financial Account Number
Individual Taxpayer Identification Number (ITIN)
Name
Other
Social Security Number (including masked or last four digits)

Please explain the other type(s) of PII that this project uses.

Date of Birth, Place of Birth, Mailing Address

Cite the authority for collecting SBU/PII/FTI (including SSN if relevant).

PII for federal tax administration - generally IRC Sections 6001 6011 or 6012
SSN for tax returns and return information - IRC section 6109

Product Information (Questions)

1 Is this PCLIA a result of a specific initiative or a process improvement?

Yes

1.1 What is the name of the Business Unit (BU) or Agency initiative?

Information Technology

2 What type of project is this (system, project, application, database, pilot/proof of concept/prototype, power platform/visualization tool)?

System

3 What Tier designation has been applied to your system? (Number)

4

4 Is this a new system?

No

4.1 Is there a previous Privacy and Civil Liberties Impact Assessment (PCLIA) for this project?

Yes

4.11 What is the previous PCLIA number?

7460

4.12 What is the previous PCLIA title (system name)?

Information Returns Processing (IRP)

4.2 You have indicated this is not a new system; explain what has or will change and why. (Expiring PCLIA, changes to the PII or use of the PII, etc.)

Significant Merging with Another System

5 Is this system considered a child system/application to another (parent) system?

Yes

5.1 Identify the parent system's approved PCLIA number.

7460

5.2 Identify the parent system's name as previously approved.

Information Returns Processing, IRP

6 Indicate what OneSDLC State is the system in (Allocation, Readiness, Execution) or indicate if you go through Information Technology's (IT) Technical Insertion Process and what stage you have progressed to.

Readiness

7 Is this a change resulting from the OneSDLC process?

No

8 Please provide the full name and acronym of the governance board or Executive Steering Committee (ESC) this system reports to.

Not Applicable, all governance activities are paused as of 4/15/2025.

9 If the system is on the As-Built-Architecture (ABA), what is the ABA ID number of the system? If this PCLIA covers multiple applications shown on the ABA, please indicate the ABA ID number(s) for each application covered separated by a comma. If the system is not in the ABA, then contact the ABA (<https://ea.web.irs.gov/aba/index.html>) for assistance.

Yes: 210629

10 Does this system disclose any PII to any third party outside the IRS?

Yes

10.1 Does the system have a process in place to account for such disclosures in compliance with IRC 6103(p)(3)(A) or Subsection c of the Privacy Act?

Yes

11 Does your project/system involve any use of artificial intelligence (AI), including virtual assistant, chat bot, and robotic process automation, as defined in Executive Order 13960 and 14110?

No

12 Does this system use cloud computing?

Yes

12.1 Please identify the Cloud Service Provider (CSP), FedRAMP Package ID, and date of FedRAMP authorization.

T Cloud, FR1801046750, April 23, 2025

12.2 Does the CSP allow auditing?

Yes

12.21 Who has access to the CSP audit data (IRS or 3rd party)?

IRS, TIGTA

12.3 Please indicate the background check level required for the CSP (None, Low, Moderate or High).

High

13 Does this system/application interact with the public?

No

14 Describe the business process allowing an individual to access or correct their information. (Due Process)

This system does not replace any individual taxpayer's right to due process, as dictated by the Internal Revenue Manual guidelines. IRS policy allows individual taxpayers whose data is in the system the opportunity to clarify or dispute negative determinations through the appeals process. However, due to the nature of this system, individuals may not receive specific notice that their information has been collected. The IRP System performs analysis on the data to identify potential compliance issues. The IRP System does not make any adjustments or assessments. The information is reviewed by IRS employees in the Wage & Investment and Small Business/Self Employed Business Units to correspond with the taxpayer, advising them of the proposed action (either a tax adjustment to an existing assessment or establishing an initial assessment for a tax period). The taxpayers are requested to concur or provide additional information. When applicable, the taxpayer is advised of their statutory appeal rights. Notice, consent and due process are provided in the tax forms instructions filed by the taxpayer, and pursuant to 5 USC.

16 Identify what role(s) the IRS and/or the contractor(s) performs; indicate what access level (to this system's PII data) each role is entitled to. (Include details about completion status and level of access of the contractor's background investigation was approved for.)

IRS: System owners (read-only), users (read-only), managers (read-only)

Contractor: Developers (read-only)

17 The Privacy Act of 1974 (5 USC § 552a(e)(3)) requires each agency that maintains a system of records, to inform each individual requested to supply information about himself or herself. Please provide the Privacy Act Statement presented by your system or indicate a Privacy Act Statement is not used and individuals are not given the opportunity to consent to the collection of their PII.

Notice, consent, and due process are provided in the tax forms instructions filed by the taxpayer, and pursuant to 5 USC.

18 How many records in the system are attributable to IRS Employees? Enter "Under 50,000", "50,000 to 100,000", "More than 100,000" or "Not Applicable".

50,000 to 100,000

19 How many records in the system are attributable to contractors? Enter "Under 5,000", "5,000 to 10,000", "More than 10,000" or "Not Applicable".

Under 5,000

20 How many records in the system are attributable to members of the public? Enter "Under 100,000", "100,000 to 1,000,000", "More than 1,000,000" or "Not applicable".

More than 1,000,000

22 How is access to SBU/PII determined and by whom?

Data access to the IRP systems and IRP storage libraries is restricted based on the principles of least privilege and separation of duties; access is granted on a need-to-know basis. IRP personnel are required to apply for access using the Business Entitlement Access Request System (BEARS). BEARS enrollment process requires that an authorized manager approve access requests on a case-by-case basis. Upon approval, IRP users are assigned user accounts based on their role(s) and responsibilities. Access to the IRP mainframe and production/development libraries and tools is restricted by Resource Access Control Facility (RACF) permissions maintained by General Support System 21 (GSS-21) International Business Machines (IBM) Mainframe Platform administrators. Each sub-application in IRP requires that a user complete a separate BEARS for the component they are requesting access to. Based on the individual's role(s) and responsibilities, access is either approved or denied, and the appropriate corresponding changes are made to the RACF access control list by the Executive Operations IBM Masterfile Security Administrators. The IBM Masterfile Security Administrators provide initial user identification and password access to the Master File. This is managed by placing users into assigned groups. Each group is assigned the minimum user rights and permissions to perform their function.

Users are restricted from changing the boundaries of their access. IRP user accounts require annual management recertification. Employees that no longer require access are required to be removed by management. Account revocation is also performed through BEARS. IRP contains read-only informational return data. By design, IRP users are restricted from inputting data into the application by Role Based Access Controls and are given the minimum set of privileges required to perform their regular and recurring work assignments.

23 Is there a data dictionary on file for this system? Note: Selecting "Yes" indicates an upload to the Attachment Section is required.

No

24 Explain any privacy and civil liberties risks related to privacy controls.

There are no privacy and civil liberties risks related to privacy controls. Data access to the IRP systems and IRP storage libraries is restricted based on the principles of least privilege and separation of duties; access is granted on a need-to-know basis.

25 Please upload all privacy risk finding documents identified for the system (Audit trail, RAFT, POA&M, Breach Plan, etc.); click "yes" to confirm upload(s) are complete.

Yes

26 Describe this system's audit trail in detail. Provide supporting documents.

Our application uses the existing mainframe and EDP platforms, which already have integration with ESAT for auditing purposes. All commands sent to the mainframe and queries/interactions with EDP are logged to ESAT already. There is no direct access to IRP data. All access is through batch files. The data viewed by authorized IRS employees is a copy of the IRMF data loaded into any number of other systems that have a user interface. These systems maintain the authentication and authorization required, including the use of audit trail information. IRP relies on the Resource Access Control Facility (RACF)/General Support System (GSS)-21 IBM Masterfile Platform to capture login information for developer access to the IRP systems.

27 Does this system use or plan to use SBU data in a non-production environment?

Yes

27.1 Please upload the Approved Email and one of the following SBU Data Use Forms, Questionnaire (F14664) or Request (F14665) or the approved Recertification (F14659). Select Yes to indicate that you will upload the Approval email and one of the SBU Data Use forms.

Yes

Interfaces

Interface Type

IRS Systems, file, or database

Agency Name

EDP

Incoming/Outgoing

Both

Transfer Method

Electronic File Transfer Utility (EFTU)

Interface Type

IRS Systems, file, or database

Agency Name

IBM Mainframe

Incoming/Outgoing

Both

Transfer Method

Electronic File Transfer Utility (EFTU)

Systems of Records Notices (SORNs)

SORN Number & Name

IRS 24.046 - Customer Account Data Engine Business Master File

Describe the IRS use and relevance of this SORN.

The Privacy Act requires Federal agencies that maintain a system of records to publish systems of records notices (SORNs) in the Federal Register for records from which information is retrieved by any personal identifier for an individual who is a US citizen or an alien lawfully admitted for permanent residence. The Privacy Act also provides for criminal penalties for intentional noncompliance.

SORN Number & Name

IRS 34.037 - Audit Trail and Security Records

Describe the IRS use and relevance of this SORN.

The Privacy Act requires Federal agencies that maintain a system of records to publish systems of records notices (SORNs) in the Federal Register for records from which information is retrieved by any personal identifier for an individual who is a US citizen or an alien lawfully admitted for permanent residence. The Privacy Act also provides for criminal penalties for intentional noncompliance.

SORN Number & Name

IRS 24.030 - Customer Account Data Engine Individual Master File

Describe the IRS use and relevance of this SORN.

The Privacy Act requires Federal agencies that maintain a system of records to publish systems of records notices (SORNs) in the Federal Register for records from which information is retrieved by any personal identifier for an individual who is a US citizen or an alien lawfully admitted for permanent residence. The Privacy Act also provides for criminal penalties for intentional noncompliance.

SORN Number & Name

IRS 22.061 - Information Return Master File

Describe the IRS use and relevance of this SORN.

The Privacy Act requires Federal agencies that maintain a system of records to publish systems of records notices (SORNs) in the Federal Register for records from which information is retrieved by any personal identifier for an individual who is a US citizen or an alien lawfully admitted for permanent residence. The Privacy Act also provides for criminal penalties for intentional noncompliance.

Records Retention

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Validation and Perfection

What is the GRS/RCS Item Number?

19-67

What type of Records is this for?

Electronic

Please provide a brief description of the chosen GRS or RCS item.

IRP data files are approved for deletion/destruction under a variety of National Archives and Records Administration-approved disposition authorities for specific batch programs and published in IRS Records Control Schedules (RCS) 19, 28, 29 and 35.

What is the disposition schedule?

IRP data files are approved for deletion/destruction under a variety of National Archives and Records Administration-approved disposition authorities for specific batch programs and published in IRS Records Control Schedules (RCS) 19, 28, 29 and 35.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Information Returns Master File (IRMF)

What is the GRS/RCS Item Number?

18-2(C)

What type of Records is this for?

Electronic

Please provide a brief description of the chosen GRS or RCS item.

Information Returns Master File (IRMF), under RCS 18 Item 2(C), maintains production data for the current and nine prior tax years.

The current tax year and five prior tax years are updated either weekly, monthly, or as needed.

What is the disposition schedule?

As a new tax year is added, the oldest tax year file is destroyed.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Underreporter Backup Withholding

What is the GRS/RCS Item Number?

19-64

What type of Records is this for?

Electronic

Please provide a brief description of the chosen GRS or RCS item.

Underreporter BacIRP data files are approved for deletion/destruction under a variety of National Archives and Records Administration-approved disposition authorities for specific batch programs and published in IRS Records Control Schedules (RCS) 19, 28, 29 and 35.

What is the disposition schedule?

IRP data files are approved for deletion/destruction under a variety of National Archives and Records Administration-approved disposition authorities for specific batch programs and published in IRS Records Control Schedules (RCS) 19, 28, 29 and 35.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Non-filer

What is the GRS/RCS Item Number?

19-69

What type of Records is this for?

Electronic

Please provide a brief description of the chosen GRS or RCS item.

IRP data files are approved for deletion/destruction under a variety of National Archives and Records Administration-approved disposition authorities for specific batch programs and published in IRS Records Control Schedules (RCS) 19, 28, 29 and 35.

What is the disposition schedule?

IRP data files are approved for deletion/destruction under a variety of National Archives and Records Administration-approved disposition authorities for specific batch programs and published in IRS Records Control Schedules (RCS) 19, 28, 29 and 35.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Underreporter

What is the GRS/RCS Item Number?

35-31

What type of Records is this for?

Electronic

Please provide a brief description of the chosen GRS or RCS item.

IRP data files are approved for deletion/destruction under a variety of National Archives and Records Administration-approved disposition authorities for specific batch programs and published in IRS Records Control Schedules (RCS) 19, 28, 29 and 35.

What is the disposition schedule?

IRP data files are approved for deletion/destruction under a variety of National Archives and Records Administration-approved disposition authorities for specific batch programs and published in IRS Records Control Schedules (RCS) 19, 28, 29 and 35.

Data Locations

What type of site is this?

System

What is the name of the System?

Mainframe (IRMF VSAM)

What is the sensitivity of the System?

Federal Tax Information (FTI)

Please provide a brief description of the System.

The IRMF Master File (IRMF) ecosystem is responsible for storing and processing information returns submitted to IRS including the following information: income, payments,

retirements contributions withholding, etc. IRMF is primarily used to match third-party reports with taxpayer filings to detect underreporting, non-filing, or fraud.

What are the incoming connections to this System?

Authoritative IRP Data via the Mainframe (IRMF VSAM and PMF VSAM) -> Informatica CDC/EFTU -> IRMF (AWS S3)

What are the outgoing connections from this System?

IRMF (AWS S3) -> Databricks Pipelines -> Databricks Bronze/Silver Layer -> Databricks Gold Layer Pipeline -> Databricks Gold Layer -> Databricks Extract Pipelines -> AWS S3 bucket -> EFTU->Mainframe-> Internal and External Systems