

Date of Approval: 06/13/2025
Questionnaire Number: 2155

Basic Information/Executive Summary

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

Information Sharing and Reporting - Analytics and Reporting (ISR-A&R)
Migration

Acronym:

ISRANR

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 ISR-A&R project sources and merges data from the following business areas - Individual Accounts, Business Accounts, Individual Return Transaction Files, Business Returns Transaction Files, Information Returns Database, Social Security Administration (SSA) Data master, Payer Master Files (Waivers and Extensions), Coverage Data Repository (Exchange Periodic Data), ACA Compliance Validation, Information Returns Master File. The primary purpose is to migrate ISR-A&R Datamart from On-prem Greenplum Managed Services to Databricks on Enterprise Data Platform (EDP) to provide uninterrupted service to the end user.

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).

Individual account, Business Account, Individual Return, Business Return, Information Return data is collected from upstream systems, processed, aggregated and persisted into the A&R Data mart in EDP-Databricks. The PII information is limited to "Dimension tables". This data is made available to downstream consumers only via Business Objects Enterprise (BOE) Reports and Tableau Dashboards. PII data is made available to restricted set of users based on Business Entitlement Access Request System (BEARS) Entitlements that need to be approved by IT and Business Management. Currently, all data will be retained as per requirements.

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

Address

Agency Sensitive Information

Citizenship or Migration Status

Credit Card Number

Document Locator Number (DLN)

Email Address

Employer Identification Number

Employment Information

Family Members

Financial Account Number

Geographical Indicators

Individual Taxpayer Identification Number (ITIN)

Name

Preparer Taxpayer Identification Number (PTIN)

Social Security Number (including masked or last four digits)

Tax ID Number

Telephone Numbers

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 the Inflation Reduction Act (IRA)?

Yes

1.1 What is the IRA Initiative Number?

4.1

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

Project - Database Migration from On-prem Greenplum to Databricks on EDP on the Treasury Cloud.

3 What Tier designation has been applied to your system?

2

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?

4395

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

System 7.3 Affordable Care Act (ACA) Data Mart

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.)

Greenplum is currently scheduled for decommissioning on Sep 30th, 2025. To maintain continuity for the business users, the database is being migrated from on Prem Greenplum to Databricks on EDP on the Treasury Cloud.

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

No

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.

Execution

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.

Enterprise Services Governance board (ESGB) and Application Development Governance Board (ADGB)

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.

210645

10 If this system discloses any PII to any third party outside the IRS, 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?

No

10.1 You have indicated that you do not have an "accounting of disclosures" process in place; please indicate a projected completion date or explain the steps taken to develop your accounting of disclosures process. Note: The Office of Disclosure should be contacted to develop this system's accounting of disclosures process.

No, reports are accessible to internal users based on their Business Entitlement Access Request System (BEARS) Entitlements.

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.

Amazon Web Services (AWS) Databricks Treasury Cloud (TC) Simple Service Storage (S3). The FedRAMP Package ID is # FR1801046750. The date of FedRAMP authorization was 03/02/2020.

12.2 Does the CSP allow auditing?

Yes

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

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

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

15 Is this system owned and/or operated by a contractor?
No

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.)
Anyone with a managerial approved Business need and managerial approved Business Entitlement Access Request System (BEARS) can be granted access. ISR-A&R Production access to PII data is limited to the ISR-A&R Database Administrators (read/write) and approved reporting applications (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.

The legal authority to collect information is Internal Revenue Code sections 6001, 6011, and 6012(a), and 6109. The principal purpose(s) for which the information is intended to be used are in the tax forms instructions filed by the taxpayer. The routine uses that may be made of the information are in the tax forms instructions filed by the taxpayer. The effects on the individual, if any, of not providing all or any part of the requested information are in the tax forms instructions filed by the taxpayer.

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".
Not Applicable

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".

Not Applicable

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?

Access to the ISR-A&R Datamart is requested via a BEARS Entitlement. Access is granted on a need-to-know basis. The BEARS Entitlement process requires that an authorized manager approve access requests on a case-by-case basis. Access approval is based on the Users role(s) and responsibilities. Users are given the minimum set of privileges required to perform their regular and recurring work assignments. They are restricted from changing the boundaries of their access without management approval. The employee's access will be terminated once they no longer require access to the Database. Deletion from the active access role is also performed through BEARS Entitlements.

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

Yes

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.

No

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

EDP Platform (Databricks) related audit logs are made available to Splunk, this includes Logon, Logoff and all DDL/DML activities. Application logs are captured within "LOG_" tables in the ANR_PIPELINE.

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

IPM

Incoming/Outgoing

Incoming (Receiving)

Transfer Method

Amazon Web Services Platform (AWS)

Interface Type

IRS Systems, file, or database

Agency Name

ISR-A&R

Incoming/Outgoing

Outgoing (Sending)

Transfer Method

Other

Other Transfer Method

Tableau Dashboards

Interface Type

IRS Systems, file, or database

Agency Name

ACV

Incoming/Outgoing

Incoming (Receiving)

Transfer Method

Amazon Web Services Platform (AWS)

Interface Type

IRS Systems, file, or database

Agency Name

ISR-A&R

Incoming/Outgoing

Outgoing (Sending)

Transfer Method

Other

Other Transfer Method

Business Objects Enterprise (BOE) Reports

Systems of Records Notices (SORNs)

SORN Number & Name

IRS 34.037 - Audit Trail and Security Records

Describe the IRS use and relevance of this SORN.

To ensure that all access is logged and reviewed to prevent unauthorized access to taxpayer information

SORN Number & Name

IRS 42.021 - Compliance Programs and Projects Files

Describe the IRS use and relevance of this SORN.

To ensure that the individual information is protected and only made available to business users on a need-to-know basis in the execution of their duties

Records Retention

What is the Record Schedule System?

General Record Schedule (GRS)

What is the retention series title?

Schedule 5.2

What is the GRS/RCS Item Number?

020

What type of Records is this for?

Electronic

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

Intermediary records. Records that meet the following conditions:

- They exist for the sole purpose of creating a subsequent record

What is the disposition schedule?

Temporary. Destroy upon creation or update of the final record, or when no longer needed for business use, whichever is later.

Data Locations

What type of site is this?

Environment

What is the name of the Environment?

Enterprise Data Platform (EDP -Treasury Cloud) - Databricks

What is the sensitivity of the Environment?

Federal Tax Information (FTI)

Please provide a brief description of the Environment.

Treasury Cloud (TC) hosts the Enterprise Data Platform that is leveraged for data integration, access, and management of the environment to empower the business users so that they can securely access the integrated tax administration data for analytical and operational use.

What are the incoming connections to this Environment?

Informatica PowerCenter (PWX) version 10.5.2 for S3

What are the outgoing connections from this Environment?

Consumer Cluster, BOE and Tableau to access data from Gold Layer.