

Date of Approval: 09/24/2025  
Questionnaire Number: 2610

## Basic Information/Executive Summary

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

Compliance Vertical - Return Review Program Legacy Component (RRPLC) data ingestion to Enterprise Data Platform (EDP) in support of Greenplum (GP) ret

Acronym:

EDP- RRPLC Data Ingestion

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.

Enterprise Data Platform (EDP) Workplace Community Cloud (WC2) is designed as a universal data hub within Amazon Web Services (AWS) Treasury Cloud (TCloud). It will contain source data from all major tax processing systems as well as supporting data necessary to conduct the IRS mission. EDP will serve as the main source of integrated taxpayer data for transactional and analytical needs. EDP will utilize AWS Databricks, Marketplace Application Programming Interface (API), Redshift, Mongo Database and third-party software such as Business Objects and Informatica. EDP TCloud will also integrate with IRS common services such as Negative Taxpayer Identification Number (NTIN). The data will only be accessible to authorized IRS personnel users that have gone through the Business Entitlement Access Request System (BEARS) approval

process and will only be provided access for their specific role. For this Release, EDP will enhance the data available on the platform by ingesting data from the Return Review Program Legacy Component (RRPLC). The data will be accessible to consumers who require data from RRPLC for reporting, fraud detection, and credit eligibility use cases.

## 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).

EDP receives sensitive data as-is from authorized IRS systems, and to that extent, the data is covered by the legal statutes already in place for collection and use of this data by IRS. EDP will inherit the mitigation/elimination processes in place by the IRS to eliminate the use of sensitive data. Use of the data will be made available only to authorized business users with its corresponding BEARS entitlements. The duration for which this data will be stored on EDP is dependent on the applicable data retention requirements established by Records and Information Management (RIM) and National Archives and Records Administration (NARA), and these requirements will be met. RRPLC data will follow the mitigation/elimination processes already in place for the EDP platform. The exact consumers of RRPLC data have not been determined. RRPLC data will be available for reporting, analytics, fraud detection, credit eligibility, and other use cases.

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

Address  
Adoption Taxpayer Identification Number  
Criminal Investigation Information  
Criminal Record  
Document Locator Number (DLN)  
Email Address  
Employer Identification Number  
Employment Information  
Family Members  
Federal Tax Information (FTI)  
Individual Taxpayer Identification Number (ITIN)  
Internet Protocol Address (IP Address)  
Name

Online Identifiers  
Other  
Personal Characteristics  
Preparer Taxpayer Identification Number (PTIN)  
Social Security Number (including masked or last four digits)  
Standard Employee Identifier (SEID)  
Tax ID Number  
Telephone Numbers

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

Tax Return and Tax Form Information; Income; Withholding; and Deduction information (Individual Master File/Business Master File), Tax Refund Amount, Tax Liability Amount; Master File Transaction (MFT) Code; Type of Tax Return Filed, IRS Forms Filed; Source of Tax Return Filing (Paper or Electronic), Tax Filing Status, Number of Dependents, Name of Dependents, Date of Birth; Tax Filing History; Employer Name, Employer Tax Identification Number, Employer Address, Employer Telephone Number, Bank Account Information, Date of Death, Device Identification, Prison/Prisoner Information, Electronic Filing Identification Number (EFIN), Tax Return Preparer Name and Employer Identification Number (EIN).

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?

No

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

Project - Ingesting data from current Return Review Program Legacy Component (RRPLC) database (listed in Executive Summary) to the platform will make the data available for Case Selection Hub (RRPLC replacement system) on the platform. The data will be accessible to Tax Examiners for reporting, fraud detection, and other use cases.

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

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?

2424

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

Case Selection Vertical - Source Data to EDP in support of Greenplum (GP) retirement

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.)

In this release the current RRPLC data will be ingested to the Enterprise Data Platform (EDP) system. The GP6 (Greenplum6) managed services platform where the RRPLC is currently hosted is being decommissioned, therefore RRPLC data will be ingested onto EDP platform.

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

Yes

5.1 Identify the parent system's approved PCLIA number.

1571

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

Enterprise Data Platform (EDP)

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.

Application Development Compliance Service Domain Governance Board (AD CD GB)

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.

The system is built under the Enterprise Data Platform (EDP) ABA #: 211416.  
We will request to update EDP's interface input on the ABA site.

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.

Not applicable because the system does not disclose any PII to any third party outside the IRS.

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.

Treasury Cloud/Workplace.gov Community Cloud FedRAMP High  
(TCloud/WC2-H) FR1801046750 09/29/2023

12.2 Does the CSP allow auditing?

Yes

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

Internal Review Service (IRS), Treasury Inspector General for Tax Administration (TIGTA) and Government Accountability Office (GAO)

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)

Individuals cannot correct their information in EDP. EDP receives data from other IRS upstream sources/systems. Those other sources/systems provide consent and due process via those systems and their related tax forms instructions, and pursuant to 5 USC.

15 Is this system owned and/or operated by a contractor?

Yes

15.1 If a contractor owns or operates the system, does the contractor use subcontractors; or do you require multiple contractors to operate, test, and/or maintain this system?

No

15.2 What PII/SBU data does the subcontractor(s) have access to?

None at this time.

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.)

Contractor/Subcontractor User: Read-only, access. Moderate background investigation level.

Contractor/Subcontractor Managers: Read-only, access. Moderate background investigation level.

Contractor/Subcontractor System Administrators: Administrator access; Moderate background investigation level.

Contractor/Subcontractor Developers: Read and write access. Moderate background investigation level.

Contractor/Subcontractor Testers: Read-only, access. Moderate background investigation level.

Only contractors/subcontractors with completed background investigation will have access to SBU/PII data on EDP. Neo4j Admin roles can view PII data.

Background investigation for all contractors is complete. Users/Administrator that need to see PII data have to submit Live Data Waiver by submitting 1) Treasury WC2 User Account \_Access or Change Request Form 2-7 (copy attached) and 2) SBU Certification of Review pdf form (copy attached).

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.

EDP receives data from other IRS upstream sources/systems. Those other sources/systems provide the Privacy Act Notice to individuals. Notice, consent and due process are provided via the IRS systems and their related tax forms instructions, 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".

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?

All contractors and employees must go through the Public Trust Clearance process before access is considered. Once cleared, access to WebApps Platform go through Business Entitlement Access Request System (BEARS). All access must be approved by the user's manager who reviews the BEARS request at the time of submission and on an annual timeframe. The system administrators/approvers will also verify group membership to ensure only the appropriate rights are granted based upon need-to-know. For non-production supporting environments users must complete the necessary Sensitive But Unclassified (live) data training, request access through the BEARS, and in some cases as outlined by the requirements set forth within the Internal Revenue Manual submit an elevated access letter that is approved by the Associate Chief Information Officer prior to granting access. The non-production environment will also routinely review access lists and verify accounts, removing ones that are no longer necessary. Every individual is reminded of their Unauthorized Access (UNAX) requirements where they are restricted to see certain taxpayer data and in many instances a third-party tool is implemented to restrict access to that data.

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 is no identified risk related to privacy controls of privacy and civil liberties.
- 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 infrastructure uses AWS standard services such as CloudWatch and CloudTrail. Logs are also ingested by the Splunk instance managed by the IRS Cyber Team. RRPLC data files will be copied to EDP AWS S3 locations. The data files will be processed using Databricks. Databricks notebooks changes are audited by Databricks natively. Audit logs go to S3, CloudWatch, and Splunk. Audit logs are available in Splunk.
- 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**

Return Review Program Legacy Component (RRPLC)

### **Incoming/Outgoing**

Incoming (Receiving)

### **Transfer Method**

Electronic File Transfer Utility (EFTU)

## **Systems of Records Notices (SORNs)**

### **SORN Number & Name**

IRS 46.002 - Criminal Investigation Management Information System and Case Files

Describe the IRS use and relevance of this SORN.

Records pertaining to possible violations of laws under the enforcement jurisdiction of the IRS, received by the IRS from other sources or developed during investigative activities, that identify or may identify criminal or civil nonconformance with Federal tax laws and other Federal laws delegated to Criminal Investigation (CI) for investigation or enforcement; information arising from investigative activities conducted by Criminal Investigation in conjunction with other Federal, state, local, or foreign law enforcement, regulatory, or intelligence agencies; personal, identification, criminal history, and other information, including information sources, pertaining to individuals identified as person(s) of interest by Special Agents assigned to the Dignitary Protection Detail; personnel and workload management information. Records include biographical, travel, communication, financial, and surveillance information. To maintain, analyze, and process sensitive investigative information that identifies or may identify criminal noncompliance with Federal tax laws and other Federal laws delegated to CI for investigation or enforcement, and that identifies or may identify the individuals connected to such activity. To establish linkages between identity theft and refund or other tax fraud schemes, and the individuals involved in such schemes, that may be used to further investigate such activity and to perfect filters that identify fraudulent returns upon filing and to facilitate tax account adjustments for taxpayer victimized by these schemes.

**SORN Number & Name**

IRS 34.037 - Audit Trail and Security Records

Describe the IRS use and relevance of this SORN.

To identify and track any unauthorized accesses to sensitive but unclassified information and potential breaches or unauthorized disclosures of such information or inappropriate use of government computers to access Internet sites for any purpose forbidden by IRS policy (e.g., gambling, playing computer games, or engaging in illegal activity), or to detect electronic communications sent using IRS systems in violation of IRS security policy.

**SORN Number & Name**

IRS 22.061 - Information Return Master File

Describe the IRS use and relevance of this SORN.

Information Returns - Individual payors and payees of various types of income for which information reporting is required (e.g., wages, dividends, interest, etc.) to administer tax accounts related to the filing of information returns.

**SORN Number & Name**

IRS 42.021 - Compliance Programs and Projects Files

Describe the IRS use and relevance of this SORN.

To track information relating to special programs and projects to identify non-compliance schemes and to select individuals involved in such schemes for enforcement actions.

**SORN Number & Name**

IRS 22.062 - Electronic Filing Records

Describe the IRS use and relevance of this SORN.

Records pertaining to individual electronic filing providers, including applications to participate in electronic filing, credit reports, reports of misconduct, law enforcement records, Device ID, and other information from investigations into suitability for participation. Records pertaining to the marketing of electronic filing, including surveys and opinions about improving electronic filing programs. RECORDS

**SORN Number & Name**

IRS 46.050 - Automated Information Analysis System

Describe the IRS use and relevance of this SORN.

To maintain, analyze, and process records and information that may identify patterns of financial transactions indicative of criminal and/or civil noncompliance with tax, money laundering, Bank Secrecy Act, and other financial laws and regulations delegated to CI for investigation or enforcement, and that identifies or may identify the individuals connected to such activity. To establish linkages between fraudulent transactions or other activities, and the individuals involved in such actions, that may be used to further investigate such activity and to perfect filters that identify information pertaining to such activity.

**SORN Number & Name**

IRS 42.017 - International Enforcement Program Information Files

Describe the IRS use and relevance of this SORN.

Listing of individual taxpayers, Taxpayer Identification Number (TIN) (e.g., Social Security Number (SSN), Employer Identification Number (EIN), or similar number assigned by IRS), summary of income expenses, financial information as to foreign operations or financial transactions, acquisition of foreign stock, controlling interest of a foreign corporation, organization or reorganization of foreign corporation examination results, information concerning potential tax liability, records pertaining to Advanced Pricing Agreements and mutual agreements.

**SORN Number & Name**

IRS 22.054 - Subsidiary Accounting Files

Describe the IRS use and relevance of this SORN.

Documents containing name, address, Taxpayer Identification Number (TIN) (e.g., Social Security Number (SSN), Employer Identification Number (EIN), or similar number assigned by the IRS), and accounting information relevant to various transactions related to unapplied credits and payments, property held by the IRS, erroneous payments, accounts transferred, funds collected for other agencies, abatements and/or assessments of tax, uncollectible accounts, and Offers-in-Compromise.

**SORN Number & Name**

IRS 24.046 - Customer Account Data Engine Business Master File

Describe the IRS use and relevance of this SORN.

Individuals who file business tax and information returns; individuals who file other information filings; and individuals operating under powers of attorney for these businesses. Tax records for each applicable tax year or period, including employment tax returns, partnership returns, excise tax returns, retirement and employee plan returns, wagering returns, estate tax returns; information returns; representative authorization information; and Device ID.

**SORN Number & Name**

IRS 22.026 - Form 1042S Index by Name of Recipient

Describe the IRS use and relevance of this SORN.

Records include taxpayer's name, address, country of residence and Taxpayer Identification Number (TIN) (e.g., Social Security Number (SSN), Employer Identification Number (EIN), or similar number assigned by the IRS), and name of withholding agent.

**SORN Number & Name**

IRS 24.030 - Customer Account Data Engine Individual Master File

Describe the IRS use and relevance of this SORN.

Tax records for each applicable tax period or year, representative authorization information (including Centralized Authorization Files (CAF)), Device ID and a code identifying taxpayers who threatened or assaulted IRS employees.

# Records Retention

What is the Record Schedule System?

General Record Schedule (GRS)

What is the retention series title?

Transitory and intermediary

What is the GRS/RCS Item Number?

5.2 item 20

What type of Records is this for?

Electronic

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

RRPLC will validate EFDS data on EDP to support need for various business units within IRS to be used for transactional and / or analytical purpose.

What is the disposition schedule?

All data meeting end of retention period requirements will be eliminated, overwritten, degaussed, and/or destroyed in accordance with National Archives and Records Administration (NARA)-approved disposition authorities for that system's data and done so in the most appropriate method based upon the type of storage media used in accordance with IRM 1.15.6.10.

What is the Record Schedule System?

General Record Schedule (GRS)

What is the retention series title?

Information Systems Security Records

What is the GRS/RCS Item Number?

3.2 Item 031

What type of Records is this for?

Electronic

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

EDP will monitor and capture records as part of the user identification and authorization process to gain access to systems (System Access Records (AU 11)). Currently, access to EDP requires corresponding BEARS entitlements.

What is the disposition schedule?

Temporary. Destroy 6 years after password is altered or user account is terminated, but longer retention is authorized if required for business use. All data meeting end of retention period requirements will be eliminated, overwritten, degaussed, and/or destroyed when business use ceases or in accordance with National Archives and Records Administration (NARA)-approved disposition authorities for that system's data and done so in the most appropriate method based upon the type of storage media used in accordance with IRM 1.15.6.10.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Tax Administration Collection - Online Payment Agreement (OPA)

What is the GRS/RCS Item Number?

28 item 158

What type of Records is this for?

Electronic

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

Direct Debit Installment Agreements (DDIA) (Form 433 Series) and related documents are records used by Compliance function taxpayer contact personnel to set up an agreement between the IRS and the taxpayer. The completed form permits the taxpayer to pay delinquent taxes through installment payments.

What is the disposition schedule?

Destroy immediately after 12 years. All data meeting end of retention period requirements will be eliminated, overwritten, degaussed, and/or destroyed in accordance with National Archives and Records Administration (NARA)-approved disposition authorities for that system's data and done so in the most appropriate method based upon the type of storage media used in accordance with IRM 1.15.6.10.

What is the Record Schedule System?

General Record Schedule (GRS)

What is the retention series title?

Information Systems Security Records

What is the GRS/RCS Item Number?

3.2 Item 030

What type of Records is this for?

Electronic

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

EDP will monitor and capture records as part of the user identification and authorization process to gain access to systems (System Access Records (AU 11)). Currently, access to EDP requires corresponding BEARS entitlements.

What is the disposition schedule?

All data meeting end of retention period requirements will be eliminated, overwritten, degaussed, and/or destroyed when business use ceases or in accordance with National Archives and Records Administration (NARA)-approved disposition authorities for that system's data and done so in the most appropriate method

based upon the type of storage media used in accordance with IRM 1.15.6.10.

## Data Locations

What type of site is this?

System

What is the name of the System?

Enterprise Data Platform (EDP)

What is the sensitivity of the System?

Federal Tax Information (FTI)

Please provide a brief description of the System.

EDP Workplace Community Cloud (WC2) is designed as a universal data hub within AWS Gov Cloud (Treasury Cloud). It will ultimately contain source data from all major tax processing systems as well as supporting data necessary to conduct the IRS mission. EDP WC2 will serve as the main source of integrated taxpayer data for analytical needs. EDP WC2 will utilize Amazon Web Services Databricks, Marketplace API, Redshift and third-party software such as Informatica, Business Objects and Tableau. Authorized IRS personnel will use analytical tools such as Business Objects and Tableau to generate reports leveraging the taxpayer data. EDP WC2 will also integrate with IRS common services such as NTIN to prevent unauthorized access to taxpayer data. EDP WC2 platform currently hosts multiple datasets - DDIA (Direct Debit Installment Agreement), CADE2 ODS (Individual Taxpayer Information), TAMS (Tax Account Management Services), MICA (Modernized Individual Custodial Accounting), BMF (Business Master File), IRMF (Individual Return Masterfile), Clean Energy (CE), and Direct File. Furthermore, new data APIs were deployed, enabling Read/Write Access to Business Taxpayer Account (BTA) and Clean Energy (Clean Vehicle Credits) users. All data on boarded onto EDP will only be accessible to authorized IRS personnel users that have gone through the Business Entitlement Access Request System (BEARS) approval process and only be provided access for their specific role.

What are the incoming connections to this System?

DDIA (Direct Debit Installment Agreement), CADE2 ODS (Individual Taxpayer Information), TAMS (Tax Account Management Services), MICA (Modernized Individual Custodial Accounting), BMF (Business Master File), IRMF (Individual Return Masterfile), Clean Energy (CE), Direct File, and Modernized Tax Return Database (MTRDB) via EFTU, Informatica IEP, or APIs.

What are the outgoing connections from this System?

DDIA (Direct Debit Installment Agreement), CADE2 ODS (Individual Taxpayer Information), TAMS (Tax Account Management Services), MICA (Modernized Individual Custodial Accounting), BMF (Business Master File), IRMF (Individual Return Masterfile), Clean Energy (CE) via EFTU, Informatica IEP, or APIs.

What type of site is this?

System

What is the name of the System?

Splunk

What is the sensitivity of the System?

Sensitive But Unclassified (SBU)

Please provide a brief description of the System.

Splunk is a Security Information and Event Management (SIEM) software solution tool composed of various dashboards that more and our Security Organization is using to aggregate and/or analyze security data for systems/applications.

What are the incoming connections to this System?

Splunk connects to our project database and operating systems.