Date of Approval: 02/07/2025
Questionnaire Number: 1691

Basic Information/Executive Summary

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

Corporate Data Initiative

Acronym:

CDI

Business Unit

Small Business and Self Employed

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 Corporate Data Initiative (CDI) is an umbrella product that provides authoritative databases, comprehensive inventory management, case processing, notice generation and reporting capabilities for candidate applications that need to improve or automate existing work processes or implement new legislative mandates. CDI is fully compliant with Federal Information Security

Modernization Act (FISMA) requirements, IRS standards and disaster recovery, allowing CDI to quickly create solutions for legacy applications that have outdated, manual, inefficient, or unsustainable processes. There are four applications under the CDI umbrella: • The Transmittal (e3210) application is an electronic version of Form 3210, Document Transmittal that is required when shipping information that contains Personally Identifiable Information (PII) data. The e3210 allows the creation, tracking, follow-up, and acknowledgement of receipt of sensitive data. • The Tax Equity and Fiscal Responsibility Act (TEFRA) application provides inventory management, with the ability to move cases from initial contact to closure. This includes various phases of responses and assessments. • The Employer Shared Responsibility Payment (ESRP) application provides inventory analysis and selection for Exam Case Selection by calculating applicable ESRP amounts and employer/employee pre-contact data. It provides inventory management, with the ability to move cases from initial contact to closure. This includes various phases of responses, recalculation, and assessments. • The Affordable Care Act Non-Filer (ACANF) application provides inventory analysis and selection for potential information return non-filers by calculating applicable failure to file and failure to furnish penalties. It provides inventory management, with the ability to move cases from initial contact to closure. This includes various phases of responses, recalculation, and assessments. CDI applications provide multiple levels of reporting for users, as well as workload reporting and up-stream reporting for all levels of management and policy analysts. TEFRA and e3210 use M365 PowerApps for the user interface. ESRP and ACANF use a custom web-based user interface. All CDI application data is stored and managed in Microsoft Structured Query Language (SQL) Server databases. Active Directory is used for user login authorization. All users must request access to the individual application subsystems and are validated before access is granted. Standard security and monitoring applications are utilized. All components are internal to the IRS. No external connections are made.

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).

Social Security Numbers (SSN) are currently required by CDI applications to assimilate data from multiple sources in the assembly of case files. Portions of the data come from IRS information returns which contain Employer Identification

Numbers (EIN) and their related employee SSN's which are needed for correlation. Additionally, portions of CDI data are sourced from the Health Coverage Tax Credit System (HCTC). CDI does not interact directly with HCTC but relies upon the SSN data to correlate the information with employer reported data.

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

Address

Email Address

Employer Identification Number

Employment Information

Federal Tax Information (FTI)

Internet Protocol Address (IP Address)

Name

Online Identifiers

Social Security Number (including masked or last four digits)

Standard Employee Identifier (SEID)

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)

1.1 Is this PCLIA a result of the Inflation Reduction Act (IRA)?

No

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

System

1.35 Is there a data dictionary for this system?

Yes

1.36 Explain in detail how PII and SBU data flow into, through and out of this system.

PII is captured in CDI by users in specific user roles with privileges to initiate limited data import and export functions during prescribed activities. For example, a limited number of CDI Employer Shared Responsibility Payment (ESRP) users initiate an annual import of processing year data from reports that are generated by the Affordable Care Act (ACA) Compliance Verification (ACV) process, but ESRP does not interact directly with ACV. The reports contain data pulled from return transaction files containing transcribed line items from employer filed information returns. The files contain data such as employer name, address, and Employer Identification Number (EIN), employee name, and social security number (SSN). Employee SEID, Name, Device, and IP address information is captured initially in CDI applications when users are added to the system, and thereafter referenced to fulfill audit log requirements for designated actions within the system. Agents are installed on application and database servers, which collect and send data to the audit logging system, SPLUNK. PII, FTI and SBU are stored in encrypted Microsoft Structured Query Language (SQL) Server databases. Access to the data is restricted to users with approved entitlements within the Business Entitlement Access Request System (BEARS) that enforce separation of duties and limit access to the minimum amount of information required to perform designated tasks.

1.4 Is this a new system?

No

1.5 Is there a Privacy and Civil Liberties Impact Assessment (PCLIA) for this system?

Yes

1.6 What is the PCLIA number?

7706

1.7 What are the changes and why?

Updating PCLIA to include IP Address in the list of PII data elements captured, as well as the use of SPLUNK for downstream audit log collection.

1.8 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.

211310

1.9 What OneSDLC State is the system in (Allocation, Readiness, Execution)?

Execution

1.95 If this system has a parent system, what is the PCLIA Number of the parent system?

Not Applicable

2.1 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? Contact Disclosure to determine if an accounting is required. Enter "Yes" or "No". If Exempt, type "Exempt".

No

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

Small Business Self-Employed (SBSE) Technology Solutions Governance Board (TGB)

3.1 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?

No

3.3 Does this system use cloud computing?

No

3.6 Does this system interact with the public through a web interface?

No

3.7 Describe the business process allowing an individual to access or correct their information.

Taxpayers receive notices and/or letters with details on the disposition of cases within CDI. They have an opportunity to dispute or correct the data in their response to the notice or letter.

4.1 Who owns and operates the system (IRS Owned and Operated, IRS Owned and Contractor Operated, Contractor Owned and Operated)?

IRS Owned and Operated

4.2 If a contractor owns or operates the system, does the contractor use subcontractors?

No

4.5 Identify the roles and their access level to the PII data. For contractors, indicate whether their background investigation is complete or not.

IRS Role: Access Level: Users Read and Write Managers Read and Write Sys. Administrators Administrator Developers Administrator

4.51 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

4.52 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

4.53 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".

More than 10,000

4.54 If records are attributable to a category not mentioned above in 4.51 through 4.53, please identify the category and the number of corresponding records to the nearest 10,000. If none, enter "Not Applicable".

Not Applicable

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

CDI utilizes Business Entitlement Access Request System (BEARS) to document approvals for access. Data access is granted on a need-to-know basis. A potential user must submit a request for access to their local management for approval. Users are not permitted access without a signed form from an authorized management official. Specific permissions (Read, Write, Modify, Delete, and/or Print) are defined on the form and set (activated) by the System Administrator prior to the user being allowed access. User privileges and user roles determine the types of data that each user has access to. Management monitors system access and removes permissions when individuals no longer require access. Users are assigned to specific modules of the application and specific roles within the modules and accounts follow the principle of least privilege which provide them the least amount of access to PII/SBU data that is required to perform their business function after receiving appropriate approval.

5.1 Please describe any privacy risks, civil liberties and/or security risks identified for the system that need to be resolved and what is the mitigation plan?

The CDI Security and Privacy Assessment Report (SAR) for FISMA24 is attached, as well as a POAM report for Issues 52888, 52908, and 52910.

5.11 Is there a Risk Assessment Form and Tool (RAFT) associated with this system on file with your organization or the IRS Risk Office.

No

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

Yes

5.3 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

Partnership Control System (PCS)

Incoming/Outgoing

Incoming (Receiving)

Agency Agreement

No

Transfer Method

Electronic File Transfer Utility (EFTU)

Interface Type

IRS Systems, file, or database

Agency Name

Affordable Care Act Compliance Validation (ACV)

Incoming/Outgoing

Incoming (Receiving)

Agency Agreement

No

Transfer Method

Other

Other Transfer Method

Manual data import initiated by end user with appropriate permissions.

Interface Type

IRS Systems, file, or database

Agency Name

Streaming Data Monitoring Tool (SDMT) Splunk

Incoming/Outgoing

Outgoing (Sending)

Agency Agreement

No

Transfer Method

Other

Other Transfer Method

Audit log file data is transferred by SDMT Splunk Universal Forwarder (UF) to the SDMT Splunk Indexer.

Systems of Records Notices (SORNs)

SORN Number & Name

IRS 42.021 - Compliance Programs and Projects Files

Describe the IRS use and relevance of this SORN.

CDI applications reference Compliance Files data for identification and management of cases meeting application-specific criteria.

SORN Number & Name

IRS 24.030 - Customer Account Data Engine Individual Master File

Describe the IRS use and relevance of this SORN.

CADE IMF data is needed to match individual employee records with employer records for determination of employer shared responsibility payment.

SORN Number & Name

IRS 34.037 - Audit Trail and Security Records

Describe the IRS use and relevance of this SORN.

CDI produces audit logs with all required auditable events based on FIPS 199 categorization and makes logs available for transfer by SDMT Splunk Universal Forwarder (UF) to the SDMT Splunk Indexer.

SORN Number & Name

IRS 24.046 - Customer Account Data Engine Business Master File

Describe the IRS use and relevance of this SORN.

CADE BMF data is needed to match employer records with individual employee records for determination of employer shared responsibility payment.

SORN Number & Name

IRS 42.001 - Examination Administrative Files

Describe the IRS use and relevance of this SORN.

CDI applications reference Examination Administrative Files data for identification and management of cases meeting application-specific criteria.

Records Retention

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Partnership Control System (PCS)

What is the GRS/RCS Item Number?

RCS 28 Item 152 C

What type of Records is this for?

Both (Paper and Electronic)

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

PCS outputs include a variety of weekly, monthly and quarterly reports. Report output types are outlined in IRM 4.29.4. The reports are used primarily by programs located within Submissions Processing Campuses, but also by field Examination and Appeals Business Units. The statute data is downloaded into the Audit Information Management System (AIMS) to facilitate the manipulation of the data.

What is the disposition schedule?

Destroy 3 years after processing year, or when no longer needed for audit or operational purposes, whichever is sooner.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Partnership Control System (PCS)

What is the GRS/RCS Item Number?

RCS 28 Item 152 B

What type of Records is this for?

Both (Paper and Electronic)

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

The Partnership Control System Data Store (PCS DS) contains relationship information among partners and partnerships along with related individual returns being examined. The data is extracted from other systems and form types.

What is the disposition schedule?

Destroy when no longer needed for audit or operational purposes whichever is sooner.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Health Coverage Tax Credit System (HCTC)

What is the GRS/RCS Item Number?

RCS 18 Item 68 B

What type of Records is this for?

Electronic

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

The Health Coverage Tax Credit System (HCTC) maintains eligibility information for all potentially eligible and eligible participants. Includes the name, address, SSN, date of birth for the participant as well as all qualified family members, the participant's monthly payment information provided by US Bank, historical data for all eligible participants and participants to include payment history and account activities as well as any necessary casework or logs.

What is the disposition schedule?

Delete/Destroy 10 years after cutoff or when no longer needed for administrative, investigative, legal, audit or other operational purposes, whichever is later.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Routine Transmittal Letters and Memoranda.

What is the GRS/RCS Item Number?

RCS 23 Item 36

What type of Records is this for?

Both (Paper and Electronic)

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

Records pertaining to the shipment and receipt of returns and documents within the examination function (includes Form 3210, Document Transmittal).

What is the disposition schedule?

Destroy when 1 year old.