

Date of Approval: 08/28/2024
Questionnaire Number: 1373

Basic Information/Executive Summary

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

Customer Account Data Engine 2

Business Unit

Taxpayer Services

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 Customer Account Data Engine 2 (CADE 2) established database(s) that house the individual taxpayer accounts, which provides IRS employees the ability to view updated account information on-line. The application performs daily maintenance to the database keeping the CADE 2 database data in sync with IMF Daily Processing (DP) taxpayer account updates on the master file. CADE 2 balances the financials on the database to the IMF and includes regular data integrity checks. The CADE 2 application includes custom code written in JAVA, Common Business-Oriented Language (COBOL), Extract Transform and Load (ETL) Informatica code, COTS product JCL, and Assembly Language Code (ALC) programming languages. The CADE 2 application code is executed on the GSS-21 mainframe and provides the following key CADE 2 functions: Daily Update (DU) processing - using identification and extraction of individual taxpayer account data changes occurring during Daily Processing (DP); Beginning of Year (BOY) and Mid-Year Conversion (MYC) changes - annual database update processing of individual taxpayer account data changes from

Individual Master File (IMF); Data Access Services (DAS) - provides responses to IMFOL command codes using predefined queries of the CADE 2 database for CFOL; Balance and Control - produce financial reports to verify database updates to financial information balances are consistent and balances with IMF data; and Data Correction Tools - The Data Correction Tools (code and processes) used to examine, troubleshoot, and, if necessary, modify production CADE 2 data. CADE 2 consists of three databases: the Active Primary Database - stores data from IMF Entity levels 1 through 3 and Tax Modules levels 1 to 4; the Active Replica Database - replica of the Active Primary Database; and the Archive Database - stores data removed from the Active Primary Database (Not used by downstream systems).

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 Personally Identifiable Information (PII) collected from the IRS forms 1040 and 1040X, and all supplemental documentation filed along with an individual's tax information, collected by upstream systems and sent to CADE 2, is used to validate an individual's taxes. The only Sensitive But Unclassified (SBU)/PII data CADE 2 uses are those necessary to assess taxes. This includes the SSN since it is the one unique identifier that taxpayers must identify themselves. CADE 2 runs daily updates with data coming from IMF. As this data is transferred and updated, it is also verified through balance and control measures. The database validates the data and ensures completeness by assessing the format and flagging errors that could affect a tax return assessment. During this processing, if any data elements are determined to be outside of the expected formatting parameters, the element is flagged for further analysis by the system. If the system is unable to correct the element, a ticket is created, and the element will be examined by appropriate personnel that can resolve the issue. Taxpayer information will continue to be processed by using IMF, with parallel validation, until CADE 2 is officially accepted by both the Chief Financial Officer and Government Accountability Office (GAO) as the authoritative data source for individual tax account data. In addition, when other internal systems that rely on CADE 2 data discover inaccurate or incomplete information due to their direct interaction with individual taxpayer, the information is resubmitted through appropriate processes and the taxpayer information is updated. In addition, the CADE 2 Database Conversion (DBC) process conducts an IMF Annual and Mid-Year Conversion

and retains the CADE 2 historical data (changes to transactions, tax modules, balances, tax return records, taxpayer records, and pending events).

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
Agency Sensitive Information
Centralized Authorization File (CAF)
Document Locator Number (DLN)
Employer Identification Number
Federal Tax Information (FTI)
Financial Account Number
Individual Taxpayer Identification Number (ITIN)
Name
Preparer Taxpayer Identification Number (PTIN)
Protected Information
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)

1.1 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.

The Personally Identifiable Information (PII) collected from the IRS forms 1040 and 1040X, and all supplemental documentation filed along with an individual's tax information, collected by upstream systems and sent to CADE 2, is used to validate an individual's taxes. The only Sensitive But Unclassified (SBU)/PII data CADE 2 uses are those necessary to assess taxes. This includes the SSN since it is the one unique identifier that taxpayers must identify themselves.

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?

6187

1.7 What are the changes and why?

PCLIA 6187 expires 8/18/2024

1.8 If the system is on the As-Built-Architecture, what is the ABA ID of the system? If this PCLIA covers multiple applications shown on the ABA, please indicate the ABA ID for each application covered separated by a comma.

210274

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

Execution

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".

Exempt

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

Strategic Development Executive Steering Committee (SDESC)

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.

The authority for processing taxpayer information is 5 U.S.C. 301 and 26 U.S.C. 7801. The purpose for sharing taxpayer information received by other IRS systems and processed by CADE 2 is to assess and distribute tax returns. Information from CADE 2 is shared with IDRS to provide data for open cases and is shared with BDA and IPM for the purpose of providing data for a new data store that will address downstream system data requirements. CADE 2 is only a repository for taxpayer data. CADE 2 does not interact directly with taxpayers like other systems regarding return transactions and authorized taxpayer representatives.

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 System administrators with administrator access, IRS developers with read and write access, contractor developers with read and write access and high background investigation level, contractor system administrators with administrator access and high background investigation level. All contractors and employees must go through the Public Trust Clearance process before access is considered. Once cleared, an IRS employee or contractor must complete the proper request forms before access to CADE 2 is obtained. All access must be approved, via the Business Entitlement Access Request System (BEARS), by the user's manager who reviews the access request at the time of submission and on an annual basis to verify the position request and if the user has a need-to-know. The system administrators/approvers will also verify group membership to ensure system rights are limited based on the employee or contractor's need-to-know to perform their official duties.

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 members of the public? Enter “Under 5,000”, “5,000 to 10,000”, “More than 10,000” or “Not applicable”.
More than 10,000

4.6 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, an IRS employee or contractor must complete the proper request forms before access to CADE 2 is obtained. All access must be approved, via the Business Entitlement Access Request System (BEARS), by the user's manager who reviews the access request at the time of submission and on an annual basis to verify the position request and if the user has a need-to-know.

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?

Risk Based Decision (RBD) 15807555 was created to address the FLID Compare Tool that does not have the negative TIN check (NTIN) capability. The tool is designed to query both the CADE 2 database and the IMF VSAM files for the same accounts and identify any differences. The FLID Compare tool returns all the data for the account (or module only if that is all that is being validated). IRS does not have the resources to implement NTIN check. The residual risk is so minimal it does not justify the cost to implement NTIN. The chance of accessing one's own data or that of a covered relationship is extremely unlikely. IRM 5.1.3.7.5.1.1 related to UNAX would be followed if data were accessed.

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
Integrated Cust Communications Environment Interactive
Processor (ICCEIP)
Incoming/Outgoing
Outgoing (Sending)
Transfer Method
Other
Other Transfer Method
Network

Interface Type
IRS Systems, file, or database
Agency Name
Enforcement Revenue Information System (ERIS)
Incoming/Outgoing
Outgoing (Sending)
Transfer Method
Other
Other Transfer Method
Network

Interface Type
IRS Systems, file, or database
Agency Name
Individual Master File (IMF)
Incoming/Outgoing
Incoming (Receiving)
Transfer Method
Other
Other Transfer Method
Network

Interface Type
IRS Systems, file, or database
Agency Name
Individual Master File Online (IMFOL)
Incoming/Outgoing
Outgoing (Sending)
Transfer Method
Other
Other Transfer Method
Network

Interface Type
IRS Systems, file, or database

Agency Name
National Account Profile (NAP)
Incoming/Outgoing
Outgoing (Sending)
Transfer Method
Other
Other Transfer Method
Network

Interface Type

IRS Systems, file, or database
Agency Name
Executive Control Program for IMF Extract (IMF701EXEC)
Incoming/Outgoing
Incoming (Receiving)
Transfer Method
Other
Other Transfer Method
Network

Interface Type

IRS Systems, file, or database
Agency Name
Integrated Production Model (IPM)
Incoming/Outgoing
Outgoing (Sending)
Transfer Method
Other
Other Transfer Method
Network

Interface Type

IRS Systems, file, or database
Agency Name
Enterprise Data Platform (EDP)
Incoming/Outgoing
Outgoing (Sending)
Transfer Method
Other
Other Transfer Method
Network

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 Customer Account Data Engine is used in tax administration for the federal government.

SORN Number & Name

IRS 34.037 - Audit Trail and Security Records

Describe the IRS use and relevance of this SORN.

The Customer Account Data Engine maintains a trail of actions taken on tax account records.

SORN Number & Name

IRS 24.030 - Customer Account Data Engine Individual Master File

Describe the IRS use and relevance of this SORN.

The Customer Account Data Engine is a modernized copy of the Individual Master File.

Records Retention

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Information Technology

What is the GRS/RCS Item Number?

Schedule 17 RCS # 17

What type of Records is this for?

Electronic

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

System Acceptability Test Documentation (SAT). Includes project folders, system test plans, predetermined results, test data, Problem Definitions, copies of supporting documentation, i.e., PRPs, Unified Work Requests, Computer Operator's Handbook, FSPs, Core Record Layouts, and End of Test Status Reports.

What is the disposition schedule?

Destroy/Delete or Overwrite when recordkeeping copy is produced.

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Information Technology

What is the GRS/RCS Item Number?

Schedule 17 RCS # 19

What type of Records is this for?

Electronic

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

Tax Processing Systems Software and Database Design

Programming Documentation Records developed and maintained for service center mainframe operating/security systems, telecommunications, transaction processing, and program transmittal software. All programs support National and field operations. Records include instructions or guidelines establishing program objective, programming requirements, documentation, design specifications, data dictionaries and related records.

What is the disposition schedule?

Destroy when no longer needed for current operations.

Data Locations

What type of site is this?

System

What is the name of the System?

GSS-21 IBM Mainframe

What is the sensitivity of the System?

Federal Tax Information (FTI)

Please provide a brief description of the System.

IBM Mainframe.

What are the incoming connections to this System?

Individual Master File (IMF) for record of tax administration.

What are the outgoing connections from this System?

Corporate Files On-Line (CFOL) for record of tax administration.