Customer Account Data Engine (CADE), Release 3.1, Milestone 4A, Version 0.3
Privacy Impact Assessment

PIA Approval Date – July 13, 2007

Requested Operational Date – Currently Operational

System Overview

The general purpose of CADE is to accept, validate, and store taxpayer and tax return information. CADE is a key component of the modernized application environment as an enabler of greatly improved processing. The Business Unit owner of CADE is the Wage and Investment (W&I) division of IRS.

Over several releases, CADE will create accurate, current, authoritative data stores as the Taxpayer Account Database (TADB) and the Tax Return Database (TRDB) in the Modernization Blueprint (2000). CADE will also construct related tax administration systems processes, incrementally replacing the IRS master files.

Data in the System

1. Generally describe the information to be used in the system in each of the following categories: Taxpayer, Employee, and Other.

Taxpayer–CADE maintains information about taxpayers for the purpose of administering the tax code. Information maintained on individuals is derived from information submitted by the taxpayer on tax forms 1040, 1040A, 1040EZ including those with schedules for credit interest and ordinary dividends and disabled credits. CADE Release 3.1 will be initialized with 1040EZ taxpayer accounts, containing five years or less of tax modules. For detailed information regarding the data elements in CADE, please consult the data dictionary found in Appendix A of this document. This section contains information on the data elements from Service Center Inputs Processing Automation System (SCIPAS), the Individual Master File (IMF), and the Enterprise Application Integration Broker (EAIB).

CADE obtains the following data from IRS systems:

1. **Service Center Inputs Processing Automation System (SCIPAS)** - PII and related data elements found in the IRS 1040, 1040A, and 1040EZ forms are received via SCIPAS which is a component of MITS-21 IBM Master File. The PII data elements transferred from SCIPAS include:

   Form 1040 U.S. Individual Income Tax Return

   a. Full name
   b. SSN
   c. Spouse's full name
   d. Spouse's SSN
   e. Address
   f. Phone number
   g. Marital status
   h. Dependents' names
   i. Dependents' SSN
   j. Occupation
   k. Spouse's occupation

l. Taxable income
m. Routing number
n. Account number
o. Designee's name
p. Designee's phone number
q. Designee's PIN
r. Paid preparer's name
s. Paid preparer's SSN or PTIN
t. Preparing firm's EIN
u. Preparing firm's phone number

Form 1040A U.S. Individual Income Tax Return

a. Full name
b. SSN
c. Spouse's full name
d. Spouse's SSN
e. Address
f. Phone number
g. Dependents' names
h. Dependents' SSN
i. Marital status
j. Taxable income
k. Routing number
l. Account number
m. Designee's name
n. Designee's phone number
o. Designee's PIN
p. Occupation
q. Spouse's occupation
r. Paid preparer's name
s. Paid preparer's PTIN or SSN
t. Preparing firm's EIN
u. Preparing firm's address
v. Preparing firm's phone number

Form 1040EZ U.S. Individual Income Tax Return

a. Full name
b. SSN
c. Spouse's full name
d. Spouse's SSN
e. Address
f. Phone number
g. Marital status
h. Taxable income
i. Routing number
j. Account number
k. Designee's name
l. Designee's phone number
m. Designee's PIN
n. Occupation

o. Spouse's occupation
p. Paid preparer's name
q. Paid preparer's PTIN or SSN
r. Preparing firm's EIN
s. Preparing firm's address
t. Preparing firm's phone number

2. **Individual Master File (IMF)** – The following IMF files supply information to CADE.
   a. ***Data Master File (DM-1)*** - The Social Security Administration (SSA) is the authoritative source for the Data Master File (DM-1) which it provides to IMF. IMF is the conduit for receiving the (DM-1) into CADE. The following data elements are included in the DM-1 file:
      - Tin
      - Tin type code
      - Citizenship
      - Byte count
      - Zeroes
      - Nap record code
      - Master file source CD
      - Data base defer CD
      - Update date (yyyymmdd)
      - Taxpayer date of birth (yyyymmdd or '00000000')
      - Gender CD
      - Name control count
      - Validation source code
      - New this quarter indicator
      - Name control

   b. ***Date of Death (DOD) File-*** SSA also provides to IMF the Date of Death (DOD) file which is sent to CADE. The following elements are included in DOD file:
      - Record code
      - SSN
      - Last name
      - Suffix
      - First name
      - Middle name
      - Very proof code
      - Date of death (mmddyyyy)
      - Date of birth (mmddyyyy)
      - State
      - Zip code
      - Zip of lump sum
      - Filler

   c. Direct Deposit Limitation File - IRS maintains the Direct Deposit Limitation file (460-07-10) of bank accounts utilized for direct deposit of Form 1040 tax refunds. The following data elements are in the system for the taxpayer:
      - Routing and transit number
      - Bank account number

- Unpostable code
- Married filing separate
- Record count
- Validity digit
- SSN
- DLN
- Tax period
- Filing status
- Cycle
- Zip code
- Refund amount

d. ***Treasury Offset Program (TOP) Debt File –*** The TOP Debt File is provided by Financial Management Services (FMS) to IMF. Each week IMF provides the TOP Debt File to CADE. The TOP Debt File includes the following data elements:
   - Taxpayer Identification Number (TIN)
   - Tax year
   - Debt Amount

e. ***IMF to CADE Initialization –*** CADE is initialized with taxpayer accounts, containing five years or less of tax modules. Data elements received during the IMF to CADE initialization include:
   - Name
   - Previous Name
   - SSN
   - TIN
   - MFT
   - Address
   - City
   - Universal Location Code
   - Area Office
   - Zip Code
   - Marital Status
   - Spouse's Name
   - Spouse's SSN
   - Year of Spouse's Death
   - Foreign Operations Post of Duty
   - Disaster Code
   - Combat Zone Code
   - Income Data
   - Federal Employee Indicator
   - Handicap Indicator
   - Preparer's TIN
   - Bank Account Number
   - Routing and Transit Number

f. ***701 Exec Return Transaction File (RTF)*** – Initialization: Taxpayer returns that have been validated and accepted by the IMF. This transfer data from the CPE to CADE is through shared DASD or tape files controlled by RACF and NCONTROL. This activity occurs during the initialization and as part of the normal processing.
    - Name
    - SSN
    - MFT
    - Address
    - City
    - Universal Location Code
    - Area Office
    - Zip Code
    - Marital Status
    - Spouse's Name
    - Spouse's SSN
    - Dependent's Name
    - Dependent's SSN
    - Dependent's Date of Birth
    - Foreign Operations Post of Duty
    - Citizenship Code
    - Income Data
    - Handicap Indicator
    - Preparer's SSN
    - Preparer's EIN
    - Preparer's telephone number
    - Third Party Designee's Name
    - Third Party Designee's Phone Number
    - Third Party Designee's TIN
    - Bank Account Number
    - Routing and Transit Number

3. **Enterprise Application Integration Broker (EAIB)** – EAIB is a collection of technologies offered through the MITS-18 infrastructure that implement service-oriented architecture (SOA) solutions. It will be used to service requests ftom Accounts Management Systems (AMS) and broker the information needed from the request to CADE R3.1 service providers. EAIB is a gateway to CADE and is not a target system; as such, it reformats the AMS request and relays it to CADE R3.1 for response. The EAIB will identify CADE specific transactions and forward the transactions to CADE.
    a. ***Get Taxpayer Address of Records Request*** - The following data elements are provided to CADE during this transaction**:**
        - Identification Number
        - Number Type Code

- Name Control
- Universal Request ID
- Timestamp
- Version number

b. ***Update Taxpayer Address of Record Request*** - The following data elements are provided to CADE during this transaction**:**
    - Identification Number
    - Number Type Code
    - Name Control
    - Update Address Type Code
    - Universal Request ID
    - Timestamp
    - DLN
    - Update address response type code
    - Version number
    - Address line 2
    - Street address
    - City
    - State
    - Zip code
    - Zip plus four code
    - Country name

In addition to the taxpayer data in the production environment, the CADE System Acceptability Test (SAT) effort uses copies of taxpayer data to complete testing activities in the development environment. Although CADE does not have a documented process to manage the controls of live data (PII) it uses the guidance provided in the IRM 2.5.16, Use of Live Data in Testing for acquiring and using live data.)

a. Must have compelling justification for using live data in test
b. Identification of where live data is located
c. Who has access to live data
d. How it is purged when no longer required (Identify artifact(s))

This live data is not scrubbed in the Development Environment. Access to LIVE DATA (for Developers) is managed by Live Data Waiver Process. CADE Production Files are identified with Generation Data Group (GDG) and retention dates in the JCL.

Taxpayer data elements provided to CADE from external sources

The only external connection which CADE receives data elements from is Financial Management Services (FMS) within Treasury. FMS will return an acknowledgement to CADE when a refund is paid or an offset occurs. This acknowledgement has the following data elements:

- Run date
- Run time
- Campus
- Agency file dsn

- Schedule number
- Control number
- File ID: IRSIND
- IM file DSN
- File type
- Cycle in file:
- Error notice payments
- Regular payments
- Grand total

Employee–CADE also maintains employee information in the form of audit logs for CADE administrative users only that record requests from CADE's system administrators, however no other personal information is used by CADE. Four classes of Audit Trails are available in the CADE system, each of which has its own categories of auditable events. The audit trails for CADE include the following areas:

- Resource Access Control Facility (RACF) Security Monitoring (at the MITS GSS 21 level)
- Landmark's "The Monitor" (TMON) for DB2 Security Monitoring (CADE related DB2 events captured by RACF)
- DB2 Transaction logging (CADE DB2 subsystem events monitored by DB2)
- CADE Subsystem logging (Application logging for CADE subsystem events)
- CADE events pass to EAIB for audit monitoring. - Event information for an EAIB Address Query and Address Update are logged in CADE's External Transaction Table. Captured information includes Unique Message Identified, Transaction Type, Timestamp and Result Code (successful or failed) data elements.

RACF Security Monitoring Events - The RACF security auditing features are enabled to ensure that audit trails are produced by the system. The audit trail allows identification of auditable events for CADE administrative users only and for the management of audit trails (logs) in a secure environment. The IRM 10.8.32 audit specification includes:

- Logon User ID which is a mainframe user account ID
- Logon Terminal ID
- Password Change Including UserID
- Password Change including Terminal ID
- File Create including file name
- File Delete including file name
- File Open including file name
- Time/Date Stamping

The data elements within the CADE audit trail do not contain any items which could be used to uniquely identify

RACF for DB2 Security Monitoring Events - The CADE DB2 database is audited for the following events:

- Access attempts denied because of inadequate authorization
- Explicit GRANT and REVOKE, assignment or change of authorization ID.

In addition, the audit logs track the time and date of the event and the user ID associated with events that fail. Furthermore, Customer Service Representative (CSR) requests for address changes to CADE does not produce employee identifiable auditable events.

Event information for an EAIB Address Query and Address Update are logged in CADE's External Transaction Table. Captured information includes Unique Message Identifier, Transaction Type, Timestamp and Result Code (successful or failed) data elements.

Other – Other refers to data in the system which is not directly associated to a taxpayer or employee.

FMS will return an acknowledgement to CADE when a refund is paid or an offset occurs. This acknowledgement has the following data elements:

- Run date
- ALC
- Run time
- Campus
- Agency file dsn
- Schedule number
- Control number
- File ID: IRSIND
- IM file DSN
- File type
- Cycle in file:
- Error notice payments
- Regular payments
- Grand total

What are the sources of the information in the system?

CADE obtains the following data from IRS systems:

1. **Service Center Inputs Processing Automation System (SCIPAS)** -PII and related data elements found in the IRS 1040, 1040A, and 1040EZ forms are received via SCIPAS) which is a component of MITS-21 IBM Master File. The PII data elements transferred from SCIPAS include:

   Form 1040 U.S. Individual Income Tax Return

   a. Full name
   b. SSN
   c. Spouse's full name
   d. Spouse's SSN
   e. Address
   f. Phone number
   g. Marital status
   h. Dependents' names
   i. Dependents' SSN
   j. Occupation
   k. Spouse's occupation
   l. Taxable income

m. Routing number
n. Account number
o. Designee's name
p. Designee's phone number
q. Designee's PIN
r. Paid preparer's name
s. Paid preparer's SSN or PTIN
t. Preparing firm's EIN
u. Preparing firm's phone number

Form 1040A U.S. Individual Income Tax Return

a. Full name
b. SSN
c. Spouse's full name
d. Spouse's SSN
e. Address
f. Phone number
g. Dependents' names
h. Dependents' SSN
i. Marital status
j. Taxable income
k. Routing number
l. Account number
m. Designee's name
n. Designee's phone number
o. Designee's PIN
p. Occupation
q. Spouse's occupation
r. Paid preparer's name
s. Paid preparer's PTIN or SSN
t. Preparing firm's EIN
u. Preparing firm's address
v. Preparing firm's phone number

Form 1040EZ U.S. Individual Income Tax Return

a. Full name
b. SSN
c. Spouse's full name
d. Spouse's SSN
e. Address
f. Phone number
g. Marital status
h. Taxable income
i. Routing number
j. Account number
k. Designee's name
l. Designee's phone number
m. Designee's PIN
n. Occupation
o. Spouse's occupation

p. Paid preparer's name
q. Paid preparer's PTIN or SSN
r. Preparing firm's EIN
s. Preparing firm's address
t. Preparing firm's phone number

2. **Individual Master File (IMF)** – The following IMF files supply information to CADE.
   c. ***Data Master File (DM-1)*** - The Social Security Administration (SSA) is the authoritative source for the Data Master File (DM-1) which it provides to IMF. IMF is the conduit for receiving the (DM-1) into CADE. The following data elements are included in the DM-1 file:
      - Tin
      - Tin type code
      - Citizenship
      - Byte count
      - Zeroes
      - Nap record code
      - Master file source code
      - Filler
      - Data base defer code
      - Update date (yyyymmdd)
      - Taxpayer date of birth (yyyymmdd or '00000000')
      - Gender code
      - Name control count
      - Validation source code
      - New this quarter indicator
      - Name control

   d. ***Date of Death (DOD) File-*** SSA also provides to IMF the Date of Death (DOD) file which is sent to CADE. The following elements are included in DOD file:
      - Record code
      - SSN
      - Last name
      - Suffix
      - First name
      - Middle name
      - Very proof code
      - Date of death (mmddyyyy)
      - Date of birth (mmddyyyy)
      - State
      - Zip code
      - Zip of lump sum
      - Filler

   e. Direct Deposit Limitation File - IRS maintains the Direct Deposit Limitation file (460-07-10) of bank accounts utilized for direct deposit of Form 1040 tax refunds. The following data elements are in the system for the taxpayer:
      - Routing and transit number
      - Bank account number

- Unpostable code
- Married filing separate
- Record count
- Validity digit
- SSN
- DLN
- Tax period
- Filing status
- Cycle
- Zip code
- Refund amount

f. ***Treasury Offset Program (TOP) Debt File –*** The TOP Debt File is provided by Financial Management Services (FMS) to IMF. Each week IMF provides the TOP Debt File to CADE. The TOP Debt File includes the following data elements:
   - Taxpayer Identification Number (TIN)
   - Tax year
   - Debt Amount

g. ***IMF to CADE Initialization –*** CADE is initialized with taxpayer accounts, containing five years or less of tax modules. Data elements received during the IMF to CADE initialization include:
   - Name
   - Previous Name
   - SSN
   - TIN
   - MFT
   - Address
   - City
   - Universal Location Code
   - Area Office
   - Zip Code
   - Marital Status
   - Spouse's Name
   - Spouse's SSN
   - Year of Spouse's Death
   - Foreign Operations Post of Duty
   - Disaster Code
   - Combat Zone Code
   - Income Data
   - Federal Employee Indicator
   - Handicap Indicator
   - Preparer's TIN
   - Bank Account Number
   - Routing and Transit Number

h. ***701 Exec Return Transaction File (RTF)*** – Initialization: Taxpayer returns that have been validated and accepted by the IMF. This transfer data from the CPE to CADE is through shared DASD or tape files controlled by RACF and NCONTROL. This activity occurs during the initialization and as part of the normal processing.
    - Name
    - SSN
    - MFT
    - Address
    - City
    - Universal Location Code
    - Area Office
    - Zip Code
    - Marital Status
    - Spouse's Name
    - Spouse's SSN
    - Dependent's Name
    - Dependent's SSN
    - Dependent's Date of Birth
    - Foreign Operations Post of Duty
    - Citizenship Code
    - Income Data
    - Handicap Indicator
    - Preparer's SSN
    - Preparer's EIN
    - Preparer's telephone number
    - Third Party Designee's Name
    - Third Party Designee's Phone Number
    - Third Party Designee's TIN
    - Bank Account Number
    - Routing and Transit Number

i. ***Communicate to CPE to IMF Annual Conversion***–Occasionally during the mid-year, IMF must make changes to the layouts of Entity and/or Tax Modules. The same layout changes must be made to the Entity and Tax Modules in CADE. Therefore, after completing the final CADE processing cycle of each year (Cycle 52/53), the data from LAFFOL will be extracted, and the LAFF portion of each record will be made available as data files for CPE. This CADE data file will be reformatted into the new layout as part of the annual IMF Conversion Run (INY 440-01) and returned to CADE via this interface.

j. ***CADE Router/Filter–***Receives one or more incoming IMF transactions from the service centers, non-service centers, and external trading partners, including FMS

acknowledgement. Each one of these transactions is processed to determine whether it should be further processed by CADE or forwarded to IMF.

3. **Enterprise Application Integration Broker (EAIB)** – EAIB is a collection of technologies offered through the MITS-18 infrastructure that implement service-oriented architecture (SOA) solutions. It will be used to service requests ftom Accounts Management Systems (AMS) and broker the information needed from the request to CADE R3.1 service providers. EAIB is a gateway to CADE and is not a target system; as such, it reformats the AMS request and relays it to CADE R3.1 for response. The EAIB will identify CADE specific transactions and forward the transactions to CADE.

    a. ***Get Taxpayer Address of Records Request*** - The following data elements are provided to CADE during this transaction**:**
        - Identification Number
        - Number Type Code
        - Name Control
        - Universal Request ID
        - Timestamp
        - Version number

    b. ***Update Taxpayer Address of Record Request*** - The following data elements are provided to CADE during this transaction**:**
        - Identification Number
        - Number Type Code
        - Name Control
        - Update Address Type Code
        - Universal Request ID
        - Timestamp
        - DLN
        - Update address response type code
        - Version number
        - Address line 2
        - Street address
        - City
        - State
        - Zip code
        - Zip plus four code
        - Country name

Taxpayer data elements provided to CADE from external sources:

The only external connection which CADE receives data elements from is Financial Management Services (FMS) within Treasury. FMS will return an acknowledgement to CADE when a refund is paid or an offset occurs. This acknowledgement has the following data elements:

- Run date
- ALC
- Run time
- Campus

- Agency file dsn
- Schedule number
- Control number
- File ID: IRSIND
- IM file DSN
- File type
- Cycle in file:
- Error notice payments
- Regular payments
- Grand total

2.a. What IRS files and databases are used?

When initialized, CADE will receive and contain individual taxpayer-related information presently contained in the following legacy files and databases:

- The Individual Master File (IMF) – The following files and data transfers are provided by the Individual Master File and are used by CADE.
    - DM-1 File from SSA
    - Date of Death File
    - Direct Deposit Limitation File
    - Treasury Offset Program Debt File
    - The Return Transaction File (RTF)
    - IMF to CADE Initialization
    - 701 Exec Return Transaction File (RTF)
    - CPE to IMF Annual Conversion
- Enterprise Application Integration Broker via AMS
    - Get Taxpayer Address Records Request
    - Update Taxpayer Address
- Service Center Inputs Processing Automation System
    - Transfers 1040, 1040A, and 1040EZ tax information from the service centers.

In addition to the yearly extract (CADE Initialization) CADE receives data form the following:

- The Individual Master File (IMF)
- Generalized Mainline Facilities (GMF) - provides records formats, does not transmit any PII.

The following CADE data files will be created and sent to CPE for processing:

- Refund Information File (RFIF)
- Statistics of Income (SOI)
- Return to CPE and IMF
- 701 Exec, Microfilm Replacement System (MRS)
- Individual Master Files (IMF)
- Microfilm Replacement System (MRS)
- Refund Timeliness Program (RTP)

- Taxpayer Account Transcripts
- LARS-format Balance and Control Data
- Weekly Obligation Balance Data
- IMF Annual Conversion
- Duplicate Direct Deposit (DDD)
- IMF Weekly Reports
- Corporate Files Online (CFOL)
- Reciprocal Accounting Control Record (RACR
- Financial Management Information Systems (FMIS) – We send report
- Return Transaction File (RTF)
- Enterprise System Management (ESM)
- Balancing Reports
- Obligation Balance Validation Reports
- CADE Initialization to IMF.
- CADE R2CPE Reports
- Accountability Acceptance Vouchers (AAV) Reports
- CPE for Discriminate Index Function (DIF) Processing
- Interim Revenue Accounting Control System (IRACS) Refund Data
- Questionable Refund Program/Refund Interest Program/Electronic Tax Administration (QRP/RIP/ETA)
- Taxpayer Address Request (TAR) – Legacy Account Formatted File (LAFF) Summary
- Interim Revenue Accounting Control System (IRACS) Recap Data
- Martinsburg Computing Center/Processing Validation Section Recap Information (ECC-MTB_PVS Recap)
- Martinsburg Computing Center/Processing Validation Section Refund Information (ECC-MTB_PVS Refund)
- Individual Return Transaction File Online (IRTFOL)
- CPE for Address Change – this information goes to the following CPE systems:
    - Enhanced Entity Index File (EEIF)
    - Key Index File (KIF)
    - Name Search File (NSF)
    - National Account Profile (NAP)
    - Address Error Report

The Business System Architecture Report (BSAR) summarizes each of the interfaces above. The interface control documents and appendices to the System Engineering Model View detail the data elements shared.

2.b What Federal Agencies are providing data for use in the system?

CADE receives the DM-1 file from IMF which contains SSA data and the TOP Debt File. However, this information is not provided directly from the agency to CADE. The only external file that CADE receives directly is the FMS acknowledgement files which are provided by Treasury as described above. The DM-1 and TOP Debt File provide current and authoritative information. The FMS is acknowledgement that the refund or offset requested by CADE was completed.

2.c. What State and Local Agencies are providing data for use in the system?

No state or local agencies provide data to CADE.

2.d. From what other third party sources will data be collected?

There are no other third party sources.

2.e. What information will be collected from the taxpayer/employee?

Information maintained on individuals is read from information submitted by the taxpayer on tax forms 1040, 1040A, 1040EZ including those with schedules for credit interest and ordinary dividends and disabled credits.

<u>Form 1040 U.S. Individual Income Tax Return</u>

a. Full name
b. SSN
c. Spouse's full name
d. Spouse's SSN
e. Address
f. Phone number
g. Marital status
h. Dependents' names
i. Dependents' SSN
j. Occupation
k. Spouse's occupation
l. Taxable income
m. Routing number
n. Account number
o. Designee's name
p. Designee's phone number
q. Designee's PIN
r. Paid preparer's name
s. Paid preparer's SSN or PTIN
t. Preparing firm's EIN
u. Preparing firm's phone number

<u>Form 1040A U.S. Individual Income Tax Return</u>

a. Full name
b. SSN
c. Spouse's full name
d. Spouse's SSN
e. Address
f. Phone number

g. Dependents' names
h. Dependents' SSN
i. Marital status
j. Taxable income
k. Routing number
l. Account number
m. Designee's name
n. Designee's phone number
o. Designee's PIN
p. Occupation
q. Spouse's occupation
r. Paid preparer's name
s. Paid preparer's PTIN or SSN
t. Preparing firm's EIN
u. Preparing firm's address
v. Preparing firm's phone number

Form 1040EZ U.S. Individual Income Tax Return

a. Full name
b. SSN
c. Spouse's full name
d. Spouse's SSN
e. Address
f. Phone number
g. Marital status
h. Taxable income
i. Routing number
j. Account number
k. Designee's name
l. Designee's phone number
m. Designee's PIN
n. Occupation
o. Spouse's occupation
p. Paid preparer's name
q. Paid preparer's PTIN or SSN
r. Preparing firm's EIN
s. Preparing firm's address
t. Preparing firm's phone number

In addition, the IRS maintains the Direct Deposit Limitation file (460-07-10) of bank accounts utilized for direct deposit of Form 1040 tax refunds. The file includes the Routing and Transit Number (RTN) of the financial institution, the specific Bank Account Number (BAN) being accessed, and the number of refunds deposited to the account during the current processing year (PY). The file also contains limited tax return information for each of those refunds. Specifically, the following data elements are included:

a. Routing and transit no.
b. Bank account number
c. Unpostable code

d. Married filing separate
e. Record count
f. Validity digit
g. SSN
h. DLN
i. MFT
j. Tax period
k. Filing status
l. Cycle
m. Zip code
n. Refund amount

3.a. How will the data collected from sources other than IRS records and the taxpayers be verified for accuracy?

Data collected from sources other than IRS records and taxpayers are verified for accuracy through the transmission integrity controls in the Connect:Direct software. Incoming data to CADE will be processed using Control-M with RACF providing identification and authentication security. Confidentiality and integrity of data processed to CADE is the responsibility of Connect:Direct. Connect: Direct will use a built-in data encryption feature to process all CADE data. CADE files created and transmitted by Control-M will not have security (confidentiality and integrity) other than identification and authentication provided by RACF.

CADE files are sent daily over the networking lines are documented using the ECC-MTB Electronic Data Transfer Questionnaire (EDTQ) memos quoted below—

- FMS to Hyattsville in Maryland (EDTQ #2002-033)

There are currently no data sharing agreements between FMS and CADE. CADE does not currently have a documented process in place for verifying the accuracy of taxpayer records from sources other that IRS records.

3.b. How will data be checked for completeness?

CADE has documented processes in place for checking completeness of data received from both internal IRS systems, as well as data received external systems. The accuracy, completeness and validity of data received, processed and transmitted through CADE are accomplished in the controls designed into the subsystem modules for Router Filter, Transaction Processing, Daily Processing, Restore to Current Processing Environment (CPE), Balance and Control, and CPE Interface.

All data files received from the original sources are checked as they enter CADE by the respective subsystem. When a subsystem identifies information that does not meet an acceptable condition(s) such as completeness of the data, taxpayers that are not in CADE, taxpayers that have been initialized/migrated into CADE but have not yet experienced processing by CADE, and taxpayers that have been initialized/migrated into CADE, and have had the tax return (TC150) posted to the modernized database, an evaluation is conducted and rejection may occur. (1.3.a provides additional information on completeness).

CADE is required to count records processed, transactions processed, pre-jouralized money amounts, and many other counts and amounts, throughout its processing in order to account for all data records and money amounts it processes. The term Balance and Control (B&C) is used to

refer to all overall mechanism for accumulating and checking these counts and amounts to ensure the overall processing integrity of CADE.

Generally, any application that processes records, transactions, or messages is required to maintain counts and amounts in the database repository known as a control block. A separate B&C application is used to run balancing formulas that check that the counts and amounts going into a step of processing equals those coming out of the step. The formulas expressing relationships among counts and amounts that must be true in order for CADE to be in balance are stored as data and processed by the B&C balance application at runtime.

EAIB service transactions being invoked by CADE will have to be examined to determine:

1. If data validation checks are performed for accuracy, completeness, validity and authenticity of information as close to the point of origin as possible.
2. Examine the information system to determine if the system employs rules for checking the valid syntax of information system inputs (e.g. character set, length, numerical range, acceptable values) to ensure that EAIB inputs match specified definitions for format and content.
3. Examine the information system to determine if the system prescreens inputs passed to interpreters to ensure EAIB content is not unintentionally interpreted as commands
4. FINALIST Address Cleansing and Duplicate Address Checking.

3.c. Is the data current? How do you know?

Yes the data in CADE is current.

EAIB Transactions have a designated response time. The time is calculated from the time CADE received a request to the time CADE sends a response back. 85% of files are processed in 3 sec—time, 98% of transactions are completed within 10 seconds.

Weekly CADE receives the DM-1 file which provides updated, current information regarding date of birth information from the SSA and is considered an authoritative source.

Each week, FMS provides the IRS with a file containing the Taxpayer Identification Number (TIN) of all the entities currently identified as having outstanding Federal debts covered under TOP. The IRS validates and sorts this file and then uses the data (File DMF-20-11) to mark refunds (Posted and RFIF) as being subject to reduction or elimination based on a TOP offset. The IRS also uses this file to trigger the conversion of a taxpayer's requested Credit Elect of an overpayment (to a subsequent obligation) to a refund, up to the amount of the TOP debt.

For CADE to properly process overpaid returns for taxpayers covered under the TOP, we must be able to identify them. Therefore, on a weekly basis, CADE will need to access CPE's TOP Debt Information File (DMF-20-11), storing the related debt amounts for all CADE TINs included in the file. (CADE will be copying information from this file. CADE will not be removing any records from the file.) CADE transmits refunds to FMS on a daily basis.

The information needed to completely answer this question was not available in any of the CADE Release 3 engineering documents.

4. Are the data elements described in detail and documented? If yes, what is the name of the document?

A Release 3.1 data dictionary provides the data elements in CADE and can be found in Appendix A. Release 3.1 DSR Data Model Interface Control Documents provides the specific data elements that make up each file transferred within a given interface.

Access to the Data

1. Who will have access to the data in the system (Users, Managers, System Administrators, Developers, Others)?

CADE RACF Group Names and descriptions are detailed in the table below:

CADE RACF Permissions

Role/RACF Group: ZCATRXG
Description: The ZCATRXG group is the authorization group associated with executing the CADE process thread from MQSeries through CICS and DB2. No userIDs will be attached to this group.
Permissions: There are no users for this group.

Role/RACF Group: OSSDB2
Description: Used by DB2 system administrators for management of DB2 systems.
Permissions: Has complete control over DB2 database on the MITS-21 platform including CADE database

Role/RACF Group: Group ZCADBAP
Description: The ZCADBAP group is the CADE Database Administrator group. This group is the owner of the tables and DDL. The group will also be responsible for defining the load.
Permissions: Delete, insert, select, and update authority.

Role/RACF Group: ZCADSUP
Description: CADE user supporting production DB2U.
Permissions: Select authority

Role/RACF Group: ZCADUSR
Description: IRS user researching and monitoring DB2U..
Permissions: Select authority

Role/RACF Group: ZCATIGTA
Description: TIGTA users for auditing purposes.
Permissions: Select authority

Role/RACF Group: Other
Description: CONTROL –M user ID (CTMCADEP) has DB2 permissions assigned to this ID for the execution of CADE jobs.
Permissions: Alter, delete, insert, select and update authority

The following employees (IRS and Contractors) who serve as privileged users have access to CADE data:

RACF Descriptions

Title/User type: Database Administrators
Description: Database Administrators design and administer the data warehousing, platform application dependencies, performance, maintenance and versioning.
RACF Group Name & Permissions: ZCADBAP permissions

Title/User type: RACF Administrators
Description: RACF Security Administrators will provide assistance in the installation, maintenance, optimization, integration, backup, and recovery of RACF. The RACF Security Administrator will also execute and support security policies and procedures.
RACF Group Name & Permissions: RI RACF Group permissions are managed by MITS-21 and are outside of the boundary of CADE.

Title/User type: System Operators
Description: System operators are responsible for day-to-day operation of CADE and associated applications, utilities, and management of both incremental and full backups of the system

Title/User type: System Programmers
Description: The principle focus of this position is the installation and maintenance of application and system software, troubleshooting, capacity planning, and performance monitoring.
RACF Group Name & Permissions: SYSADMIN (OSSDB2) permissions

Title/User type: System Developers
Description: The CADE SAT effort uses copies taxpayer data to complete testing activities
RACF Group Name & Permissions: ZCADBAP and ZCADSUP permissions

Title/User type: TIGTA Auditors
Description: Treasury Inspector General auditors
RACF Group Name & Permissions: ZCATIGTA permissions

Title/User type: Contract Employees
Description: The system programmers and developers are frequently contract employees
RACF Group Name & Permissions: Following a MBI security investigation, contract employees are granted the same permissions as IRS employees in the same capacity.

Title/User type: Taxpayers
Description: None
RACF Group Name & Permissions: None

Title/User type: Customer Service Representatives
Description: Customer Service Representative (CSR) initiates a request for taxpayer address information through the Accounts Management Services (AMS) and processes the addresses using the Integrated Data Retrieval System (IDRS). The CRS will process addresses using the IDRS. AMS through the Broker EAIB provides user interface services to select, search, view, request, print and update taxpayer address information in CADE

Title/User type: Other
Description: The Detroit Computing Center (ECC-DET) serves as the security auditors for Martinsburg Computing Center (ECC-MTB). ECC-DET monitors and reviews the activities of CADE DBAs, RACF Administrators, and Systems Programmers within the CADE data sharing group and report to ECC-MTB Security any violations or questionable activities. There remain legacy risks around auditing for CADE.

2. How is access to the data by a user determined? Are criteria, procedures, controls, and responsibilities regarding access documented?

The OL5081 is used to document access requests, modifications, and terminations for all types of users, including system administrators, and test accounts. Users requesting access to an IRS system must do so through the OL5081 form. Users are required to complete an OL5081, Information System User Registration/Change Request Form, which list mandatory rules for users of IRS information and information systems. The manager or designated official signs the OL5081 and it is submitted to the security staff or administrator, who assigns a user ID and an initial password. Before access is granted, the user is required to digitally sign OL5081, acknowledging his/her security responsibilities when using the system.

3. Will users have access to all data on the system or will the user's access be restricted? Explain.

Access to CADE is granted as described above using the OL5081 process. Access is restricted to administrative users on a need-to-know basis. The table below details the permissions for the specific levels of administrators for CADE. The RACF grants the permissions and determines the levels of access. The z/OS version 1.7 Security Server/RACF is used on the mainframe, utilizing the RACF for DB2 security features for access controls and authorizations. This access granted to the administrators will be done through their RACF privileges. These privileges are determined by the ECC-MTB security administrators based on the activities that CADE requires. Access to these privileges requires that the user have on file an IRS 5081 form and an IRS 104 form, both of which require a manager's approval and signature. Defining these privileges is the responsibility of ECC-MTB and is not controlled by CADE. Group authorizations are used wherever possible while capturing individual accountability, as required. The RACF profiles provide role-based access, and separation of duties to ensure that users can only access data as required by their jobs. By the nature of the mainframe environment administrative users have access to taxpayer data.

The following DB2 special privileges are available to users with a SELECT profile for DB2U providing role based access:

DB2 Privileges

Role: SYSADMIN (OSSDB2)
Permissions: Has complete control over the CADE database.

Role: Job Scheduler CONTROLM
Permissions: Alter, delete, insert, select and update authority

Role: Group ZCADBAP
Permissions: Delete, insert, select, and update authority.

Role: Groups ZCADSUP, ZCADUSR, ZCATIGTA,
Permissions: Select authority

Role: Other
Permissions: CONTROL–M user ID (CTMCADEP) has DB2 permissions assigned to this ID for the execution of CADE jobs. Alter, delete, insert, select and update authority

4. What controls are in place to prevent the misuse (e.g. browsing) of data by those having access?

Some CADE DBAs, as defined in the Enterprise Data Managemetn Office (EDMO) Database Administrator (DBA) Roles and Responsibilities Memorandum of Understand, do have the ability to logon to the ECC-MTB network remotely and connect to CADE resources through the normal logon procedures for ECC-MTB.

Some activities by the administrative users require that the users get a block of records instead of an individual taxpayer record. In this instance the DB2 logs would only reflect the first and last row of the record block accessed and would not show which of the records within the block the user actually viewed. Also, for CADE Release 3.1 this information is collected, but ECC-MTB does not currently have the capability to run reports against the DB2 logs. These are known problems which are currently being researched for resolutions. Until this is resolved CADE will not have any way of accessing previously unavailable data on administrative users.

However, there are some controls in place, including all CADE developers must take Live Data training on the proper handling of live data. This complements the required security awareness training for all employees and contractors.

5.a Do other systems share data or have access to data in this system? If yes, explain.

CADE is part of an integrated, enterprise-wide Tax Administration system and is the single authoritative source of Tax Return and Tax Account data for the taxpayer records that are part of CADE. CADE will transmit data to the following IRS systems:

- Refund Information File (RFIF)
- Questionable Refund Program/Refund Interest Program/Electronic Tax Administration (QRP/RIP/ETA)
- Duplicate Direct Deposit (DDD)
- Statistics of Income (SOI)
- 701 Exec, Microfilm Replacement System (MRS)
- IMF Weekly Reports
- Return to CPE and IMF
- Taxpayer Address Request (TAR) – Legacy Account Formatted File (LAFF) Summary
- Corporate Files Online (CFOL)
- Interim Revenue Accounting Control System (IRACS) Recap Data
- Interim Revenue Accounting Control System (IRACS) Refund Data
- Financial Management Information Systems (FMIS)
- Reciprocal Accounting Control Record (RACR)
- Martinsburg Computing Center/Processing Validation Section Recap Information (ECC-MTB_PVS Recap)
- Martinsburg Computing Center/Processing Validation Section Refund Information (ECC-MTB_PVS Refund)

- Individual Master Files (IMF)
- Electronic Certification System (ECS)
- Microfilm Replacement System (MRS)
- Individual Return Transaction File Online (IRTFOL)
- Return Transaction File (RTF)
- Refund Timeliness Program (RTP)
- Enterprise System Management (ESM)
- Taxpayer Account Transcripts
- Send to Current Processing Environment (CPE)
- LARS-format Balance and Control Data
- Balancing Reports
- Security Administration System (SAS) Reports
- Service Center Input Processing Automation System (SCIPAS) Reports
- Accountability Acceptance Vouchers (AAV) Reports
- Obligation Balance Validation Reports
- Weekly Obligation Balance Data
- CADE Initialization to IMF.
- IMF Annual Conversion
- CPE for Address Change – this information goes to the following CPE systems:
    - Enhanced Entity Index File (EEIF)
    - Key Index File (KIF)
    - Name Search File (NSF)
    - National Account Profile (NAP)
    - Address Error Report
    - CADE R2CPE Reports
- CPE for Discriminate Index Function (DIF) Processing
- Enterprise Application Integration Broker (EAIB)
- Application Messaging and Data Access Service (AMDAS).

CADE will transmit data to the following external systems:

- Social Security Administration Self Employment Data – Data elements provided to SSA include:
- US Census Bureau Economic Data
- Refund data to the Treasury Department and the Chief Information Officer (CIO),
- Refund data to FMS,

The Business System Architecture Report BSAR summarizes these interfaces, which are detailed in the interface control documents and appendices to the System Engineering Model View.

5.b. Who will be responsible for protecting the privacy rights of the taxpayers and employees affected by the interface?

The IRS Wage & Investment (W&I) Designated Approving Authority (DAA) is responsible for authorizing the CADE to operate, and for protecting the privacy rights of the taxpayers affected by the interface. The external interfaces are governed by Interconnection Security Agreements (ISAs). These documents provide specific details about how information is to be transferred and protected. CADE sensitive but unclassified documents are maintained in accordance with Internal Revenue Manual (IRM) 11.3. The following ISAs and Memoranda of Understanding (MOU) with the external partners are in development and maintained by Modernized Information Technology Systems (MITS) and Information Technology Security Engineering (ITSE) organization:

- Interconnection Security Agreement between IRS and FMS in Support of FMS Enterprise Systems (November 2006)
- Interconnection Security Agreement between IRS and Social Security Administration (November 2006) – Draft.

Additionally, the following individuals are also responsible for protecting the rights of the taxpayers:

- Treasury FMS - Assistant Commission, Information Resources
- SSA - Director of the Division Annual Wage Reporting Branch (DAWRB).

The information needed to completely answer this question was not available in any of the CADE Release 3 engineering documents. This requires further research and will be documented as part of CADE Release 3.1 Milestone 4a activities.

6.a. Will other agencies share data or have access to data in this system (International, Federal, State, Local, Other)?

The following Federal agencies receive data from CADE:

- Refund data to the Treasury Department and the Chief Information Officer (CIO),
- Refund data to FMS,
- Schedule SE and C data to Census, and
- Transcript data will be available to Government Accountability Office (GAO).
- SSA Self-Employment (SE) Initial Records file that contains taxpayer's SE income data and the other file is the SSA Control file (460-17-33) which contains the counts of SSA SE data record Balance and Control (B&C) information.

6.b. How will the data be used by the agency?

Data from CADE will be used by the receiving divisions of the agencies for the following:

- FMS will use refund information transmitted by CADE to do refund processing and disbursement.
- The Chief Financial Officer (FMS Reports) will use data from the CADE reports to take corrective actions on (among other things misdirected Electronic Fund Transfer (ETF) refunds. These reports will be part of CADE's process to meet the Joint Financial

Management Improvement Program (JFMIP) requirements for detailed disbursement confirmation.

- Taxpayer Account Transcripts reports taxpayer information to allow auditors to review the processing of CADE. This is intended for the use of GAO auditors.
- The file provided to census extracts the previously formatted records from LAFFOL for taxpayers that have filed tax return with Schedules SE or C and write them into the Census Data Records file for demographic and statistical purposes.
- The SSA uses the information to verify self-employment information.

6.c. Who is responsible for assuring proper use of the data?

The IRS W&I DAA, the FMS Assistant Commission, Information Resources and the SSA Director of DAWRB will be responsible for the proper use of that data at the agency. The Office of The data exchanges between CADE and external interfaces are governed by Interface Control Documents (ICDs), Interconnection Security Agreements (ISAs) and/or Memorandums of Understanding (MOUs) depending on the method of transfers and whether the interface is internal or external to the IRS.

6.d. How will the system ensure that agencies only get the information they are entitled to under IRC 6103?

CADE is required to ensure that external agencies only receive data entitled under IRC 6103. Each interface within IRS is documented using an Interface Control Document (ICD). These documents contain an overview of the purpose and scope of the interface, interface requirements, boundary conditions, interface processing, data mapping, and issues associated with the interface. Each ICD is the product of an extensive logical and physical review of the interface. The ICDs, including the data elements to be shared, are peer reviewed, commented on, and eventually approved my responsible parties. This process provides an opportunity to vet each data element against the requirements in IRC 6103. However, this could not be validated as being a standard practice for CADE. Currently there is no other process in CADE to ensure that agencies (Census, SSA, and Treasury) only get the information they are entitled to under IRC 6103. Furthermore, there is no process in place to ensure that CADE meets the requirements of M-01-05.

The IRS is currently exploring the possibility of requiring data sharing requests to accompany the ICDs as appendices. These data sharing requests shall include the authoritative names of the data sharing participants, names of systems where shared data will be stored, a detailed description of the purpose for the request, a description of each of the data categories being requested, relevance of each data category including duration and retention as specified in the retention schedules appropriate to the data in this ICD. These new requirements are being developed in response to OMB M-01-05 *Guidance on Inter-Agency Sharing of Personal Data* and in recognition of agency responsibilities associated with the sharing of IRC 6103 data as which is protected by the Privacy Act. However, there are currently no IRS programs which exist to execute and deliver these over-arching privacy requirements.

Attributes of the Data

1. Is the use of the data both relevant and necessary to the purpose for which the system is being designed?

Each interface within IRS is documented using an Interface Control Document (ICD). These documents contain and overview of the purpose and scope of the interface, the interface requirements, boundary conditions, interface processing, data mapping, and issues associated with the interface. Each ICD is the product of an extensive logical and physical review of the interface. The ICDs, including the data elements to be shared, are peer reviewed, commented on, and eventually approved my responsible parties.

The IRS is currently exploring the possibility of requiring data sharing requests to accompany the ICDs as appendices. These data sharing requests shall include the authoritative names of the data sharing participants, names of systems where shared data will be stored, a detailed description of the purpose for the request, a description of each of the data categories being requested, relevance of each data category including duration and retention as specified in the retention schedules appropriate to the data in this ICD. These new requirements are being developed in response to OMB M-01-05 *Guidance on Inter-Agency Sharing of Personal Data* and in recognition of agency responsibilities associated with the sharing of IRC 6103 data as which is protected by the Privacy Act. However, there are currently no IRS programs which exist to execute and deliver these over-arching privacy requirements.

2.a. Will the system derive new data or create previously unavailable data about an individual through aggregation from the information collected?

CADE will derive new data and create previously unavailable data about an individual through aggregation from the information collected from internal and external interfaces.

2.b. Will the new data be placed in the individual's record (taxpayer or employee)?

Yes, new data may be placed in the individual's record through the EAIB Update Taxpayer Address of Record. the Customer Service Representative (CSR) initiates a request for taxpayer address information through the Accounts Management Services (AMS) and processes the addresses using the Integrated Data Retrieval System (IDRS). The CSR will process addresses using the IDRS. AMS through the Broker EAIB provides user interface services to select, search, view, request, print and update taxpayer address information in CADE. Customer Service Representative can place new data in the individual's record through the EAIB Update Taxpayer Address Record request to CADE. This service is responsible for processing Update Taxpayer Address of Record Request and performs validation of the message for Release 3.1. The primary responsibilities of the EAIB include:

- Validate the request.
- Send the message to validate and determine address posting requirement service for posting the address change to the database.
- Perform balance and control on all incoming request and outgoing response. This includes all successful and failed responses.
- Send the response back to the EAIB.

.

2.c. Can the system make determinations about taxpayers or employees that would not be possible without the new data?

No, CADE does not have the ability to make independent determinations about taxpayers or employees. Determinations are made by other IRS systems which rely upon the data in CADE.

The new data in the system is address data which is not used to make any determinations.

2.d. How will the data be verified for relevance and accuracy?

CADE is required to count records processed, transactions processed, pre-jouralized money amounts, and many other counts and amounts, throughout its processing in order to account for all data records and money amounts it processes. The term Balance and Control (B&C) is used to refer to all overall mechanism for accumulating and checking these counts and amounts to ensure the overall processing integrity of CADE.

Generally, any application that processes records, transactions, or messages is required to maintain counts and amounts in the database repository known as a control block. A separate B&C application is used to run balancing formulas that check that the counts and amounts going into a step of processing equals those coming out of the step. The formulas expressing relationships among counts and amounts that must be true in order for CADE to be in balance are stored as data and processed by the B&C balance application at runtime.

EAIB service transactions being invoked by CADE will have to be examined to determine:

5. If data validation checks are performed for accuracy, completeness, validity and authenticity of information as close to the point of origin as possible.
6. Examine the information system to determine if the system employs rules for checking the valid syntax of information system inputs (e.g. character set, length, numerical range, acceptable values) to ensure that EAIB inputs match specified definitions for format and content.
7. Examine the information system to determine if the system prescreens inputs passed to interpreters to ensure EAIB content is not unintentionally interpreted as commands
8. FINALIST Address Cleansing and Duplicate Address Checking.

3.a If the data is being consolidated, what controls are in place to protect the data and prevent unauthorized access? Explain.

No data is consolidated in Release 3.1 .

3.b If processes are being consolidated, are the proper controls remaining in place to protect the data and prevent unauthorized access? Explain.

No processes are being consolidated as part of CADE Release 3.1 .

4. How will the data be retrieved? Can it be retrieved by personal identifier? If yes, explain.

CADE data is retrievable by the taxpayer identification number (social security number or employer identification number), document locator numbers and alphabetically by name.

Describes whether the Identification number is a Social Security Number (SSN), Individual Taxpayer Identification Number (ITIN), Employer Identification Number (EIN), or Adoption Taxpayer Identification Number (ATIN) etc.

What are the potential effects on the due process rights of taxpayers and employees of:

a. consolidation and linkage of files and systems;

b. derivation of data;

c. accelerated information processing and decision making;

d. use of new technologies;

In Milestone 4A, CADE Release 3.1 does not have a documented process to address due process rights of the taxpayer and employee as it relates to consolidation and linkage of files and systems; derivation of data; accelerated information processing and decision making; or the use of new technologies. The interconnection documents for CADE have determined:

a. Consolidation and linkage of files and systems – No data is being consolidated and not processes are being consolidated.

b. Derivation of data – CADE contains taxpayer data, including Social Security Number, name, address, and financial information. This data could be used to identify and locate the taxpayer and obtain income and tax information. Browsing or misuse of this information would be a significant privacy issue.

c. Accelerated information processing and decision making – Data will be timely and accurate and thus will result in improved service to and about the taxpayer. Daily processing will prevent inaccurate information being disseminated to taxpayers and IRS employees. CADE data is not new data and thus while it may improve overall tax processing it does not process new data for additional decision making.

d. Use of new technologies –

The purpose of the CADE EAIB interface is to manage transactions received from the EAIB and for replies returned to the EAIB. Only two types of online EAIB requests are supported by CADE under this release; A Get Taxpayer Address of Record request and an Update Taxpayer Address of Record request. Other transaction request types will return an error notification to the EAIB. The EAIB will identify the CADE transactions and forward the transactions to CADE. CADE EAIB Transactions are processed online through application requests forwarded from the Accounts Management Services (AMS) application.

How are the effects to be mitigated?

CADE does not directly make determinations on taxpayers. Determinations will be made by other systems. In Milestone 4A, CADE Release 3.1 does not have a documented process to address due process rights of the taxpayer and employee, therefore at this stage we are unable to identify how the effects are mitigated.

Maintenance of Administrative Controls

1.a. Explain how the system and its use will ensure equitable treatment of taxpayers and employees.

The IT Security Rules of Behavior (supplied by Missions Assurance) are applied to everyone accessing IRS systems (as referenced in Appendix B: Rules of Behavior).

No taxpayer or group of taxpayers is treated differently from any other taxpayer of group of taxpayers. All tax returns must adhere to the same standardized and published XML schemas and

business rules for the same return type. Schemas validate data element formats and ranges on all forms and schedules. Business rules enforce relationships between data, forms, and schedules. All returns must adhere to the same rules. There are no exceptions.

No employee or group of employees is treated differently by the system than any other employee or group of employee. All employees must have a user ID (SEID), a password, and an assigned role. All employee transactions are audited and written to the audit logs. There are no exceptions.

Procedures are in place for taxpayers that allow them to have appeal rights for taxes they feel are unjustly levied against them. In addition, taxpayers may pursue any tax problems stemming from the perceived misuse of or unauthorized access of their tax data by seeking resolution via the Taxpayers Advocate's office. Employees may be treated differently however, procedures exist that provide protections for the employee so they might not be subject to any unwarranted disciplinary actions. Form 10420, Security Incident Report and Form 11377, Inadvertent Taxpayer Data Access provide protections for those IRS employees who accessed a taxpayer's data which had not been assigned to them as a result of a keying error or inadvertent access.

CADE does possess the capability to treat individuals or groups of taxpayers differently based upon pre-established criteria. For instance, taxpayers located in an eligible disaster area are given Disaster Account Designation (DAD) processing for transaction codes TC971, AC087, 688.

1.b. If the system is operated in more than one site, how will consistent use of the system be maintained at all sites.

The system is centralized with no other sites, however CADE data is maintained through data backups. CADE data will be processes according to business rules which apply to all taxpayers.

1.c. Explain any possibility of disparate treatment of individuals or groups.

CADE does possess the capability to treat individuals or groups of taxpayers differently based upon pre-established criteria. For instance, taxpayers located in an eligible disaster area are given Disaster Account Designation (DAD) processing for transaction codes TC971, AC087, 688. Additionally, CADE possesses indicators which identify individuals serving in a combat zone, individuals serving at an overseas posting, or individuals with a handicap (blindness), which can impact the application of standard business rules. This disparate treatment is supportable by the U.S. tax code and standard IRS processing procedures.

2.a. What are the retention periods of data in this system?

The retention periods for data in the system starts at the end of the calendar year or month in which the records were created. If the records are arranged by fiscal year, list year, program year, or some standardized time period, then the retention period begins with the close of that period in which the records were created.

The z/OS Catalog system automates the Tape Backup procedures for Archival. Computer Associates backup utilities (CA-1) provides archive to automated tape libraries(ATL). Storage The silos provides archiving for Connect: Direct data.

Retention dates on production files are specified by ECC-MTB, negotiated with the Business Unit (W&I) and documented in Interface Control Documents (ICDs) for CADE. Tape expiration is managed by the automated tape library.

The following retention periods are followed for CADE:

- Data Control and Accounting Records – Destroy one year after the end of the processing year, including the following:
  - All records that form a part of the audit trail of data flow into, through, and out of ECC-MTB processing systems.
  - Ledgers and other documents pertaining to the reconciliation of the general ledger accounts in the service centers with the money balances on CADE and the master files maintained on magnetic tape at ECC-MTB.
  - Card files, tickler files and other types of files used to record action taken and control workflow.
- CADE Data Archive –This data will be retained until the next archive, it is expected that archives will take place annually.
  - CADE will archive unneeded data from the production environment. This data can be restored to the Production Support environment, if needed to support GAO audits. This data is Balance and Control, Send to MQ Status, and Statistics data. Archived data will not include taxpayer data.
- CADE Data Backups (final updated tape(s) for each calendar year) – Scratch after six months. Magnetic tape files containing current records for all taxpayers. Contains the balance, status, and transactions applicable to the individual accounts during a specific tax period. This includes returns filed, amendments to returns, assessments, debit and credit transactions.
- CADE Data Backups (all other weekly and daily backups) – Scratch after successful completion of third update cycle. Magnetic tape file containing current records for all taxpayers. Contains the balance, status, and transactions applicable to the individual accounts during a specific tax period. This includes returns filed, amendments to returns, assessments, debit and credit transactions.

Retention periods for these files are controlled by IRM 1.15.19-1, Records Control Schedule for Martinsburg Computing Center, Item 32.Work Files (Job No. NC1–58–76–8, Item 32). Work Files include: Interim Processing Media; Control Media; Print/Edit Media; Program Media; Special Project Media; Test Media; Checkpoint Media; Other Agency Media; and Unclassified Media. Retention period for these files is defined as “RELEASE for re-use when no longer needed in accordance with IRM Section 2800.”

2.b. What are the procedures for eliminating the data at the end of the retention period? Where are the procedures documented?

Files going from CADE to other systems follow the retention dates specified in the ICDs. Retention practices are managed by the system and follow applicable IRS IRMs depending upon the data type. Files within CADE are Generation Data Groups (GDGs) and follow retention rates in the JCL. Files stay on the mainframe MITS 21 until the operating system detects the date of the file. The Storage Management System (SMS) uses CA-1 from Computer Associates which writes the tapes to ATL silos where they are stored for the duration of their retention period. Disposal is managed by ATL, which detects end of life cycle. At the end of the retention period the tapes are recycled. Retention dates on production files are specified by the ECC-MTB which are negotiated with the Business Unit and documented.

A significant privacy concern does arise from the data in the development environment. Within the development partition, there is no enforced retention period for files and the data is never purged.

As a result there are no checks to ensure that live taxpayer data is not misused or retained beyond its intended retention period.

Procedures for eliminating data at the end of retention period - Records are maintained in accordance with Records Disposition Handbooks, IRM 1.15.2.12.1 through IRM 1.15.2.12.50. The following procedures are in effect at Martinsburg to dispose of data:

- ECC-MTB SOP No. 2.2.8-26 (Rev. 17), Degaussing Media.
- IRM 2.2.8, Magnetic Media Management.
- ECC-MTB SOP No. 2.2.8-29 (Rev. 21), Magnetic Media Rehabilitation.
- ECC-MTB Security Handbook Issuance No. 306, Destruction of Data.
- IRM 2.1.10, Information Systems Security.
- IRM 2.2.8, Magnetic Media Management.

CADE's media sanitization and disposal are considered a common control and the procedures in place for CADE are those in place at ECC-MTB. IRS follows disk sanitization procedures for destruction of discarded media. IRM 2.7.4, Management of Magnetic Media (Purging of SBU Data and Destruction of Computer Media) provides those procedures used for sanitizing electronic media for reuse (e.g., overwriting) and for controlled storage, handling, or destruction of spoiled media or media that cannot be effectively sanitized for reuse (e.g., degaussing). The responsibilities for management and employees for the care, cleaning, rehabilitation, storage, shipment, receipt, inspection, repair, destruction and security of all magnetic media is addressed. Coverage includes all round reels, cartridges, cassettes, removable disks, optical disks, hard drives, etc. IRM 2.7.4 also discusses duties, responsibilities, and procedures expected of all IRS sites that own, control, maintain, receive, ship, transmit, or inventory magnetic media. CADE's offsite tapes are not encrypted in transit.

2.c. While the data is retained in the system, what are the requirements for determining if the data is still sufficiently accurate, relevant, timely, and complete to ensure fairness in making determinations?

CADE does not make any independent determinations. CADE will be the custodian of the taxpayer and tax return information that is available to assist the IRS in Tax Administration. To help ensure that data in CADE remains sufficiently accurate and timely accounts are updated for each transaction and returns. In addition, payments must be processed daily, and refunds must be provided to FMS daily. To ensure completeness of the data CADE uses a balancing and reconciliation process.

3.a Is the system using technologies in ways that the IRS has not previously employed (e.g. Caller-ID)?

CADE is not employing any new technologies in Release 3.1.

3.b How does the use of this technology affect taxpayer/employee privacy?

CADE is not employing any new technologies in Release 3.1.

4.a Will this system provide the capability to identify, locate, and monitor individuals? If yes, explain.

Yes, CADE maintains tax records for certain classes of taxpayers and therefore can be used to identify and locate individuals. The inclusion of data elements such as address, zip code, and country code supports the business function of CADE and the IRS by ensuring that a taxpayer could be contacted if additional information, corrections, or correspondence were needed

regarding information submitted by the taxpayer. The monitoring of tax information provided by the taxpayer supports the business need to research cases of potential tax fraud.

4.b. Will this system provide the capability to identify, locate, and monitor groups of people? If yes, explain.

Yes, CADE will provide the capability to identify, locate, and monitor groups of people because it maintains tax records for certain classes of taxpayers.

4.c. What controls will be used to prevent unauthorized monitoring?

The following controls help prevent unauthorized monitoring:

- TMON for DB2 logs are maintained for audit purposes to monitoring a users access to the system which is necessary for
- Intrusion detection capabilities are employed to detect unauthorized monitoring
- Clear separation of duties between the RACF Administrators, DBAs, Operations Personnel, and the Security Auditor. These positions require complete separation of duties, for example a RACF Administrator can only be the Security Auditor if that RACF Administrator works in a different computing center and has only Auditor privileges at the computing center being audited. Please reference the Access to Data section of the PIA for tables detailing the separation of duties.
- Other access to the RACF system-level audit information is restricted to those people with a strict need-to-know, such as the Security Auditor and the Treasury Inspector General for Tax Administration (TIGTA).
- The audit information is protected using a RACF profile.
- Privacy training is required for all personnel, IRS and contractor.
- All administrators are subject to current security policies regarding disclosure of information.
- All employees must have a user ID (SEID), a password, and an assigned role when using the EUP. In using the EUP, all employee transactions are audited and written to the SAAS.
- Disclosure of returns and return information may be made only as provided by (1) 26 U.S.C. 3406, and (2) 26 U.S.C 6103.

5.a Under which Systems of Record Notice (SORN) does the system operate? Provide number and name.

CADE is currently covered by Treasury/IRS 24.030, CADE Individual Master File (IMF), as well Business Master File (BMF)-Treasury/IRS 24.046 and Treasury/IRS 34.047 IRS Audit Trail & Security Records System SORNs.

5.b. If the system is being modified, will the SORN require amendment or revision? Explain.

IRS Office of Disclosure has determined that CADE will not require a new System of Records Notice (SORN).

View other PIAs on IRS.gov