

Enterprise Informatica Platform (EIP) – Privacy Impact Assessment

PIA Approval Date – March 28, 2011

System Overview

The Enterprise Informatica Platform (EIP) project was developed to support multiple projects by providing a tool suite to extract, transform, and load data from the client's data source to the client's target/destination. This specifically requires the creation of development, test (staging), and production environments available to any project. The creation of EIP is a part of the Internal Revenue Service (IRS) Data Access Strategy for creating a secure, shared infrastructure and tool suite of common data services. EIP Release 1.0 will support the Customer Account Data Engine – 2 (CADE–2). Future releases of EIP will support other projects, as needed, to support the IRS Strategic Objective: Modernize the IRS through its People, Processes, and Technology. By providing a required server infrastructure component, IRS will reduce the deployment time of applications, and therefore increase the overall business value of projects from each business organization.

Systems of Records Notice (SORN):

- IRS 24.030--Customer Account Data Engine Individual Master File
- IRS 24.046--Customer Account Data Engine Business Master File
- IRS 34.037--Audit Trail and Security Records
- IRS 42.021--Compliance Programs and Project Files

Data in the System

1. Describe the information (data elements and fields) available in the system in the following categories:

A. Taxpayer – Within the staging area the following data elements are maintained:

- Taxpayer name
- Social Security Number (SSN)
- Taxpayer Identification Number (TIN)
- Taxpayer address
- Date of birth
- Any other data on the 1040 series tax forms
- Telephone number
- Income information
- Document Locator Number (DLN)
- Type of return filed (e.g., 1040, 1040A, 1040EZ)
- Source of filing (paper or electronic)
- Tax filing status
- Number of dependents
- Employer name
- Federal Employer Identification Number (FEIN)
- Employer address
- Business Master File (BMF) Entities; Tax Modules; Entity Transactions; Tax Return Transactions; Tax Return Statuses
- Transcribed sections of Business Tax Returns
- Wage Related Documents from the Social Security Administration (SSA)
- Name Control; Date of Death; Birth Data; Citizenship Indicator
- Employee Plan Master File Entities and Tax Modules

- Information Return Documents and Payer Documents
- Individual Master File (IMF) Entities; Tax Modules; Entity Transactions; Tax Return Transactions; Tax Return Statuses
- Examination Inventory and controls
- Transcribed sections of Individual Tax Returns: Form 1040
- Form 1099–MISC Income
- SS–8 Inventory and Case Development
- Earned Income Credit (EIC)
- Calculated variables based on IMF and BMF tax returns as received from the Compliance Data Environment (CDE) System
- Allowable Living Expenses – Collection Financial Standards used to determine a taxpayer's ability to pay a delinquent tax liability.
- Employment Tax Schedule R – Allocation Schedule for Aggregate Form Filers
- CP2100 Backup Withholding (BWH) "B" Notices
- Withholding Compliance Program (WHC) data

B. Audit Trail Information – Audit trail information is maintained based on the defined auditable events for the EIP application. This information includes all taxpayer information as stored on 1040 series tax forms, including name, SSN, TIN, address, date of birth, etc.

2. Describe/identify which data elements are obtained from files, databases, individuals, or any other sources.

A. IRS – Currently the only projects utilizing EIP as part of Release 1 will be CADE–2, Electronic Fraud Detection System (EFDS), and Integrated Production Model (IPM). Data elements from CADE–2 include:

- Taxpayer name
- SSN
- TIN
- Taxpayer address
- Date of birth
- All related 1040 series tax form data

Data elements from EFDS include:

- Taxpayer name
- TIN
- Taxpayer address
- Telephone number
- SSN
- Income information
- DLN
- Type of return filed (e.g., 1040, 1040A, 1040EZ)
- Source of filing (paper or electronic)
- Tax filing status
- Number of dependents
- Employer name
- FEIN
- Employer address

Data elements from IPM include:

- BMF entities
- Tax modules
- Entity transactions
- Tax return transactions
- Tax return statuses
- Transcribed sections of business tax returns
- Wage related documents from SSA
- Name Control; Date of Death; Birth Data; Citizenship Indicator
- Employee plan master file entities and tax modules
- Information return documents and payer documents
- IMF entities; tax Modules; entity Transactions; tax Return transactions; tax return statuses
- Examination inventory and controls
- Transcribed sections of individual tax returns: Form 1040
- Form 1099–MISC income
- SS–8 inventory and case development
- EIC
- Calculated variables based on IMF and BMF tax returns as received from the CDE system
- Allowable living expenses – collection financial standards used to determine a taxpayer's ability to pay a delinquent tax liability
- Employment tax schedule R – allocation schedule for aggregate form filers
- CP2100 BWH "B" notices
- WHC program data

B. Taxpayer – No data is directly obtained from taxpayers. This data is obtained directly from the individual projects. Data elements from these projects include:

- Taxpayer name
- SSN
- TIN
- Taxpayer address
- Date of birth
- All related 1040 series tax form data

3. Is each data item required for the business purpose of the system? Explain.

Yes. The purpose of EIP is to support multiple projects by providing a mechanism to extract, transform, and load data from the client's data source to the client's target/destination. Specifically, these data elements are required for the individual projects to process information. However, the data is only being stored in the temporary staging area within the EIP application.

4. How will each data item be verified for accuracy, timeliness, and completeness?

The data stored within the temporary staging area is verified from the mainframe through a series of applied business rules which are performed by the individual projects. The individual projects are responsible for all data accuracy and completeness. EIP only generates a log error if the data transfer fails.

5. Is there another source for the data? Explain how that source is or is not used.

No. There is no other source of data.

6. Generally, how will data be retrieved by the user?

Data is not directly retrieved by users, but can be accessed by authorized client developers/ analysts for that particular project through logging into the Functional Integration Test (FIT) Plus, Development, and Test environments of the EIP application. There is no user interaction while the data is being moved in the production environment. EIP administrators only have access to related log information for the individual projects.

7. Is the data retrievable by a personal identifier such as name, SSN, or other unique identifier?

Yes. While data is retrievable by personal identifiers within the FIT Plus, Development, and Test environments, it is up to the individual projects to determine the extent that a client developer/ analyst can retrieve the data. There is no user interaction while the data is being moved in the production environment.

Access to the Data

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

There is no user interaction/access while the data is being moved in the production environment. All user interaction is within the FIT Plus, Development, and Test environments. Within these environments, client developers/analysts for their particular project have access to the data within the temporary staging area. EIP administrators do not have access to project data within the temporary staging area. Users do not maintain any other type of access to EIP. The Technical Account Manager (TAM) is a contractor who is granted access to EIP for troubleshooting and technical support to individual projects.

9. How is access to the data by a user determined and by whom?

Access is determined through role-based access control. The EIP Technical Lead ensures that only those authorized administrators by the particular project have access to the temporary staging area. Each project ensures that access to their particular project is determined through the On-Line 5081 (OL5081) process for granting access. All access must be approved by the individual's appropriate manager.

10. Do other IRS systems provide, receive, or share data in the system? If YES, list the system(s) and describe which data is shared.

Yes. The data that resides in the temporary staging area is derived from each individual projects. Currently the only projects utilizing EIP as part of Release 1 will be Customer Account Data Engine (CADE)-2, Electronic Fraud Detection System (EFDS), and Integrated Production Model (IPM).

11. Have the IRS systems described in Item 10 received an approved Security Certification and Privacy Impact Assessment?

Customer Account Data Engine (CADE)-2

- CADE-2 is undergoing its initial security assessment and authorization (SA&A)
- Privacy Impact Assessment (PIA) – March 8, 2011

Electronic Fraud Detection System (EFDS)

- Security Assessment and Authorization (SA&A) – June 20, 2008
- Privacy Impact Assessment (PIA) – December 17, 2010

Integrated Production Model (IPM)

- Security Assessment and Authorization (SA&A) – August 4, 2008
- Privacy Impact Assessment (PIA) – August 17, 2010

12. Will other agencies provide, receive, or share data in any form with this system?

No. No other agencies will provide, receive, or share data in any form with EIP.

Administrative Controls of Data

13. What are the procedures for eliminating the data at the end of the retention period?

The Enterprise Informatica Platform (EIP) is non-recordkeeping. EIP is a set of applications that provides IRS employees with a secure, shared infrastructure for extracting and transporting data. The extracted data is stored in a temporary staging area. A script is needed to identify what is to be scrubbed (within 24 hours) after each successful transformation, but a records control schedule is not required.

14. Will this system use technology in a new way?

No. EIP will not use technology in a new way.

15. Will this system be used to identify or locate individuals or groups? If so, describe the business purpose for this capability.

No. EIP will not be used to identify or locate individuals or groups.

16. Will this system provide the capability to monitor individuals or groups? If yes, describe the business purpose for this capability and the controls established to prevent unauthorized monitoring.

No. EIP will not be used to monitor individuals or groups.

17. Can use of the system allow IRS to treat taxpayers, employees, or others, differently?

No. Use of EIP does not allow for the IRS to treat taxpayers, employees, or others differently.

18. Does the system ensure "due process" by allowing affected parties to respond to any negative determination, prior to final action?

Not Applicable. EIP is solely used for infrastructure for various client projects. It is up to these client projects to ensure due process is in place.

19. If the system is web-based, does it use persistent cookies or other tracking devices to identify web visitors?

No. EIP is not web-based, therefore does not use persistent cookies or other tracking devices to identify web users.

[View other PIAs on IRS.gov](#)