

Data Strategy Implementation Program (DSI) – Privacy Impact Assessment

PIA Approval Date – July 5, 2007

Requested Operational Date – Sept. 30, 2009

System Overview

The purpose of Data Strategy Implementation (DSI) Program is to implement IRS Enterprise Data Strategies that define Enterprise Data Architecture, define the current and future data requirements, close performance gap, build-out enterprise data warehouse capability, eliminate data duplication within the enterprise, provide schemes for information sharing, support interoperability among information systems, enhance FEA Data Architecture and support IRS business goals. The current stove-piped distributed environment extracts the same data of varying quality from the tax master files, tax return files and case systems to a variety of platforms, in different formats, in different locations to support analytic, reporting and statistical requirements. The solution to these is to build out an Enterprise Return Retrieval service, the Third Part Data Stores, Metadata Repository, and Data Integration Services to support IRS business unit's data needs. In this budget request, DSI will improve electronic access to the Bureau, consistent with some of the President's Management Agenda goals, the Treasury's mission, and IRS' strategic goals and objectives. DSI will reduce deployment time for software applications and therefore increase the overall business value of software development projects by improving data management. The main benefit of this solution is to provide best-in-class retrieval, process and storage of taxpayer data, which improves quality, accuracy, and lead-time-to-response.

System of Records Numbers

- IRS/Treasury 70.001 Individual Income Tax Returns, Statistics of Income
- IRS/ Treasury 44.001 Appeals Case Files
- IRS/ Treasury 42.027 Data on Taxpayers Filings on Foreign Holdings
- IRS/ Treasury 34.037 IRS Audit Trail and Security Records
- IRS/ Treasury 26.020 Taxpayer Delinquency Investigation (TDI) Files
- IRS/ Treasury 26.019 Taxpayer Delinquent Account (TDA) Files
- IRS/ Treasury 26.011 Litigation Case Files
- IRS/ Treasury 24.030 Individual CADE Master File
- IRS/ Treasury 24.046 Business CADE Master File
- IRS/ Treasury 34.037 IRS Audit Trail and Security Records System
- IRS/ Treasury 24.029 Individual Account Number File (IANF)
- IRS/ Treasury 22.062 Electronic Filing Records
- IRS/ Treasury 22.034 Individual Return Files, Adjustments and Miscellaneous Documents
- IRS/ Treasury 00.334 Third Party Reprisal Records
- IRS/ Treasury 00.333 Third Party Contact Records

Data in the System

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

- Taxpayer: The system will use taxpayer account, taxpayer return, information return and case data.
- Employee: The system will use employee type, skills, classification number, and general status data.

- Audit Trail Information (including employee log-in info): Data is strictly protected from unauthorized access. IRS audit trail protection products for the data are being built in.
- Other: The system will use Internal Management budget, cost, travel, procurement, and asset data.

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

- DSI will collect data from existing IRS systems and databases (IMF, IRTF, BMF, BRTF, GOALS, TRBD, IRMF, PMF, CADE, CAM, FPC, RC, HRConnect, IFS). DSI will use taxpayer account, taxpayer return, and case data from these existing IRS systems and databases.
- The taxpayer information is extracted from the existing IRS systems, files, and databases:
 - Individual Master File (IMF)
 - Individual Return Transaction File (IRTF)
 - Business Master File (BMF)
 - Business Return Transaction File (BRTF)
 - Treasury Financial Government Service (FMS) Government Online Accounting Link System (GOALS) II
 - Tax Return Database (TRDB)
 - Information Returns Master File Processing (IRMF)
 - Payer Master File (PMF)
 - Customer Account Data Engine (CADE)
 - CFO Accounts Receivable Management System (CAM)
 - Filing and Payment Compliance (FPC)
 - HRConnect
 - Integrated Financial System (IFS)
- From HRConnect, DSI will collect employee IDs for audit log of SQL queries.
- Treasury Financial Management Services (FMS) provides disbursement schedule confirmation information through Government Online Accounting Link System (GOALS).
- State and Local Agencies do not provide data for use in the system.
- Data is not collected from other third-party sources in DSI Releases 1-4.

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

Yes. The reports generated by the system will be provided to business owners for analysis in order to meet IRS mission goals.

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

FMS GOALS disbursement confirmation data is matched to IRS disbursement schedule data. Unmatched records are flagged for research.

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

No. Once this system is operational, it will be the primary source for this data. The system is temporarily loading data from IRS CDW system. It will load from other data sources with IRS systems.

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

DSI will retrieve data from existing IRS systems, files, and databases. The Relational Database Management System (RDBMS) table/view access control capabilities will be used to restrict user access to taxpayer/ employee data. Access to details of taxpayer transactions in DSI is restricted

to authorized personnel using batch reporting capabilities and RDBMS table/views containing taxpayer identified data is logged to create an audit trail.

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

Yes. Data is retrievable by taxpayer SSN and all other identifiers available in the contributing systems.

Access to the Data

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

- Business analysts/researchers and planning/operations managers
- Designated system and database administrators, analysts, and programmers performing system maintenance and support.
- Users, Business Owners.
- Contractors are allowed to access to data only if they have been through IRS security process and approved for access to certain data.

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

The system complies with C-2 Level and Enterprise Architecture (EA) requirements. Business owners set permissions, and monitor who has access and assure access is removed when no longer appropriate, per Security requirements.

Access is granted via the Online 5081, which is authenticated by the user's manager.

10. Do other IRS systems provide, receive, or share data in the system?

Yes. The system will share taxpayer account, return and information return data with IRS Modernization, Vision and Strategy (MV&S) and other projects.

DSI will collect data from existing IRS systems and databases (IMF, IRTF, BMF, BRTF, GOALS, TRBD, IRMF, PMF, CADE, CAM, FPC, RC, HRConnect, IFS). DSI will use taxpayer account, taxpayer return, and case data from these existing IRS systems and databases.

The taxpayer information is extracted from the existing IRS systems, files, and databases:

- Individual Master File (IMF)
- Individual Return Transaction File (IRTF)
- Business Master File (BMF)
- Business Return Transaction File (BRTF)
- Treasury Financial Government Service (FMS) Government Online Accounting Link System (GOALS) II
- Tax Return Database (TRDB)
- Information Returns Master File Processing (IRMF)
- Payer Master File (PMF)
- Customer Account Data Engine (CADE)
- CFO Accounts Receivable Management System (CAM)
- Filing and Payment Compliance (FPC)
- HRConnect
- Integrated Financial System (IFS)

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

Yes.

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

No. External organizations will not have an automated interface that allows them to share data or have access to data in this system. External organizations will receive outputs within the system from the business units following existing disclosure procedures/ processes.

Administrative Controls of Data

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

The system owner is working with Records Retention Officer to establish the proper records control & retention periods for this system. This information will be provided when this is determined.

DSI records are maintained in accordance with Records Disposition Handbook, IRM 1.15.2.1 through IRM 1.15.2.31. Input records are disposed of after publication of the statistics, except for input records for high income taxpayers for tax years prior to 1967 which are disposed of after 40 years. Output records are retained as long as they are needed. Additionally foreign tax records are generally kept for 3 year and then disposed of according to the IRMs stated above.

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

No.

15. Will this system be used to identify or locate individuals or groups?

No. DSI is not used to contact, locate, or monitor taxpayers. Only taxpayer information required for decision analytics, agency performance measurement, and management, and management reporting is extracted from operational systems and stored in DSI.

16. Will this system provide the capability to monitor individuals or groups?

No. DSI is not used to monitor an individual or groups of people.

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

No.

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

Yes. It is a tool that contributes to IRS employees' ability to make considered determinations affecting taxpayers. Still, taxpayers are entitled to due process, i.e., review and appeal rights, after such determinations are made.

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

No. The system is not Web-based.

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