

Electronic Tax Administration Research and Analysis System (ETARAS) – Privacy Impact Assessment

PIA Approval Date – Mar. 23, 2010

System Overview:

The Electronic Tax Administration Research and Analysis System (ETARAS) captures processing information from electronically filed returns for research and analysis. The system contains data on electronically transmitted individual (1040 series) tax returns, copies of Third Party Data Store (TPDS) tables, copies of the refund files transmitted to Financial Management Service (FMS), as well as information on 940, 941, 1041, 1065, 990, 720, 2290, 8849 and 1120 series returns.

Systems of Records Notice (SORN):

- IRS 22.062--Electronic Filing Records
- IRS 34.037--Treasury/IRS Audit Trail and Security Records System

Data in the System

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

- A. Taxpayer – ETARAS was created to give e-file program owners, managers and executives information about electronic filing. To that end, the system is populated with taxpayer data. Currently, that data includes information on electronically transmitted individual (1040 series) returns, copies of tables from the Third Party Data Store (TPDS), and copies of the refund files transmitted to Financial Management System (FMS). ETARAS now also has files that include balance due and payment information, as well as information on 940 series returns, 1041 series returns, 1120 series returns and 1065 series returns.
- B. Employee – ETARAS is not designed to store information on IRS employees except information on the employees who access ETARAS.
- C. Audit Trail – Audit information saved in the Application, Security and System logs are collected and removed from the servers daily. The logs are reviewed using IRS established Security Procedures and if no security issues are found, stored in a dedicated security repository server for six years. These logs include; Audit account logon events Success, Failure Audit account management Success, Failure Audit directory service access Failure Audit logon events Success, Failure Audit object access Failure Audit policy change Success, Failure Audit privilege use Failure Audit system events Success, Failure.

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

- A. IRS – ETARAS receives extracts of files from other IRS systems described in question 10 below.
- B. Employee – The only employee information stored on the system is information relative to employee log-ins.

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

Yes. Each data item is required for the business purpose of ETARAS, which is to access and view the rejected returns information in order to identify problems, like programming errors or customer misinformation, and resolve them.

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

ETARAS consolidates its data into accepted and rejected databases. It is this consolidation that gives the data its integrity. The data that ETARAS stores is unique in its format and composition. However, there are other systems, such as the E-File Report Web Page, that have summarized data, which is used to validate the accuracy, timeliness and completeness of the data. ETARAS also serves as a system of validation to these other systems, such as the E-File Reports Web Page. The ELF system produces other reports for other systems that also serve as tools for verification of the ETARAS data. Having two systems with similar data, one detail and one summary, allows us to validate one set of data against the other. Validity of the data cannot be done, since this is done at the mainframe level.

The data that ETARAS receives is the product of output reports from different major IRS systems (ELF, 94XML, MeF, EFS and ERS). These systems produce reports that become the source of the ETARAS data as well as other systems throughout the IRS. Data validation is conducted by these systems and not at the ETARAS level. These major IRS systems either accept or reject data based on accuracy and integrity of the data provided to the IRS by efile transmitters and providers. The report that ETARAS and the E-File Reports Web Site receives are identical, except that the data that ETARAS receives is detailed (one row of data for each tax return efiled to the IRS) and the data received by the Efile Reports is summarized data. Since the data that ETARAS receives is detail data, its total data must match the total summarized data received by the E-File Reports Web Site.

Analysts conduct manual checks to balance the data on these two systems. A set of coded runs has been developed to do these balancing (we have been doing these activity for more than 10 years). If an out of balance situation is found, we know that one of the systems has loaded data in an incorrect manner and procedures are in place to find where the mistake occurred and how to fix it.

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

Several systems store some of the same components of data that ETARAS stores. However, there is no single system that stores all the data that ETARAS stores on a single box. ETARAS Does not use or connect to any of these system. These system receive some but not all data from the same sources ETARAS receives its data. Different systems have different needs, so therefore, the data is not exactly the same. ETARAS being a research system, attempt to capture as much data as possible to support all the research needs of our customers. The ability of ETARAS to provide electronic return reject code information by Software Identification (SID) Numbers is unique to ETARAS. A SID number is a unique number that the IRS provides to the External Partner (Software Developers) to identify their software, which they used to transmit electronic returns, to the IRS. ETARAS is the only system nationwide with this capability.

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

Access to ETARAS is limited to e-file analysts who have a need to conduct research about various electronic filing initiatives. In some circumstances, access is given to other research functions in the IRS once a need for information to conduct research has been established. Access to the system is requested through the Online (OL) 5081 process. A signed OL5081 form is required for all access requests. Currently, no field research units have access to ETARAS but it is not unusual for them to request access or data. In addition to the users who have access to the data, users at the Enterprise Computing Centers (ECC) in Martinsburg, West Virginia and Memphis, Tennessee also have access to the systems to load the ETARAS raw data.

The ETARAS level of access is determined by the need-to-know of the user. Users who are only given access to those directories in which the EFTU files are to be uploaded.

The majority of the users retrieve the data using Structured Query Language (SQL) software. SQL allows different types of access (Read, Write, Create, etc.) within individual databases. One user in particular, uses a different software package to retrieve data from ETARAS. This IRS employee uses their IRS workstation (Windows 2005) and has the software installed on their computer. The software used is an approved IRS software. All desktop software used by users is IRS approved software. No third-party software can be used. Except for the database administrator (DBA), all users receive READ ONLY Access to the system.

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

Yes. Data is retrievable using personal identifiers (SSN, TIN, EFIN, ETIN, Software ID). There are several reasons for retrieving the data using the unique taxpayer identifiable field. The primary usage is to be able to track taxpayer-filing patterns throughout the years. Using an SSN, it is possible to track the e-filing pattern of a taxpayer from one year to another. It can be identified if a taxpayer has filed electronically using an online software, a tax practitioner or if they have not filed an electronic return at all.

Access to the Data

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

Role: Analysts

Permission: Research

Role: System Administrator

Permission: Access for system maintenance only

Role: Database Administrator

Permission: Access for system maintenance only

Note: Contractors do not have access to the application.

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

Access to ETARAS is determined by the user's need to know. The primary SA/DBA of the system makes the determination as to who will receive access to the system and the level or permission that the user will receive.

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.

ETARAS receives the following data from the following systems:

- Modernization & Information Technology Services (MITS)-24 – Electronic File Transfer Utility (EFTU) is comprised of a set of servers and agents whose job it is to monitor and control file transfer protocol (FTP) processing activity to and from production application servers. EFTU is designed to enhance the security of system-to-system file transfers between IRS internal

systems, and to automate operations, audit logging and error handling. The EFTU provides both Computing Center personnel and field personnel with a common reference point for monitoring their FTP.

ETARAS receives the following data via the EFTU:

- Electronic Filing System (ELF) 1544 Data – ELF1544 data is transmitted through FTP from the Martinsburg and Tennessee Computing Centers via the Centers' Unisys 4800 system. The FTP process is included at the end of the ELF15 run ECL script.
- Third Party Data Store (TPDS) Data – TPDS data is transmitted to ETARAS and replaces the trading partner data received from the Applicant Database (ADB).
- Financial Management Service (FMS_ Data – ETARAS receives refund data from 22 files that the IRS uses to ship information to FMS. The FTP process is initiated by ETA employees accessing data on a Martinsburg Computing Center Mainframe.
- Balance Due Payments data – This data is transmitted through FTP to ETARAS.
- Forms 940 and 941 Data – Forms 940 and 941 data is transmitted to ETARAS through an FTP process run on Tennessee Computing Center's ECL after the termination of the Magnetic Tape (MGT) Executive Control Language (ECL). The MGT ECLs transmit the TRDB and reject data files to the ETARAS system. ETARAS also receives information from the 94XML process on a special file designed and built for ETARAS.
- 1041 Data – The EFS ECL's transmit files meeting ETARAS' specifications.
- Form 1065 Data – The EFS ECL's transmit files meeting ETARAS' specifications.
- Form 1120 Data – The MeF extract transmit files meeting ETARAS' specifications.

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

Yes.

MITS –24 General Support System (GSS) (UNIX Consolidated Platform)

- Authority To Operate (ATO) – September 28, 2007, expires September 28, 2010
- Privacy Impact Assessment (PIA) – February 16, 2007, expires February 16, 2010

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

There have been some instances in which IRS has received a request from the Treasury Inspector General for Tax Administration (TITGA) and the Government Accountability Office (GAO) for ETARAS data. In these instances, the data has been provided to these agencies in the format of an extract. The type of data provided to these agencies does change depending on the type of request made. These requests are ad-hoc. These agencies must go through the IRS process of established procedures to receive IRS data.

Administrative Controls of Data

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

A request for records disposition authority for ETARAS and associated records is currently being drafted with the assistance of the IRS Records and Information Management (RIM) Program Office. When approved by the National Archives and Records Administration (NARA), disposition instructions for ETARAS inputs, system data, outputs and system documentation will be published under IRM 1.15.29 Records Control Schedule for Tax Administration – Wage and Investment Records, item number to be determined.

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

ETARAS does not use technology in a new way. Any change to the operating system and/or SQL is coordinated with IRS MITS personnel and it must follow IRS approved guidelines.

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

No, the system will not be used to identify or locate individuals or groups. However, ETARAS does have the ability to identify individual taxpayers by their SSN and monitor the activities of that SSN. The database does not allow for identification of the taxpayer name or taxpayer's address as linked to the SSN. It does, however, allow for tracking by ZIP Code. The system will also allow for tracking of return and accuracy volumes of software developers, electronic return originator, and electronic return transmitters. The control to prevent unauthorized monitoring is in restricting access. When complete, the query audit log will also note if any unauthorized monitoring has occurred. All control and review of audit trails is conducted by MITS Security Personnel and not by the ETARAS application.

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.

ETARAS does have the ability to identify individual taxpayers by their SSN and monitor the activities of that SSN. The database does not allow for identification of the taxpayer name or taxpayer's address as linked to the SSN. It does, however, allow for tracking by ZIP Code. The system will also allow for tracking of return and accuracy volumes of software developers, electronic return originator, and electronic return transmitters. The control to prevent unauthorized monitoring is in restricting access. When complete, the query audit log will also tell us if any unauthorized monitoring has occurred. All control and review of audit trails is conducted by MITS Security Personnel and not by the ETARAS application.

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

ETARAS is a research system only. Use of this system does not allow the IRS to treat taxpayers, employees, or others differently. The system is not used for compliance purposes.

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

Not applicable. This does not apply to ETARAS, since the data storage on ETARAS is used for statistical purposes and not for enforcement of any IRS regulations.

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

Not applicable. ETARAS is not web-based.

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