

# Electronic Filing of 94x XML Returns (94x XML) – Privacy Impact Assessment

PIA Approval Date – Feb. 26, 2010

## **System Overview:**

The Electronic Filing of 94x XML Returns (94x XML) is an Internal Revenue Service (IRS) Wage and Investment (W&I) application that processes Form 94x returns received electronically from preparers for business taxpayers. The system receives Employment/Unemployment Tax Returns (Forms 940, 941, and 944) and associated electronic payments from external trading partners through the Electronic Management System (EMS) interface in extensible mark-up language (XML) format. The application then validates the tax return data received and processes (e.g., extracts and reformats) the data for use with other IRS systems.

## **Systems of Records Notices (SORN):**

- IRS 22.062 – Electronic Filing Records
- IRS 24.046 – Business Master File
- IRS 34.037 – 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 – the following taxpayer personally identifiable information (PII) is contained in the system:
- Employer Identification Number (EIN)
  - Employer Name
  - Trade Association/Organization's Name, if any
  - Business address

The following is contained in the system for validation purposes:

- Electronic File (e-File) Personal Identification Number (PIN)
- Taxpayer Information Number (TIN)
- Employer Identification Number (EIN)
- Bank routing numbers
- Bank account number
- EIN and Name control (first four letters of the last name )
- Reporting Agent e-File signatures

B. Employee:

- Username
- Password

- C. Audit Trail Information – 94x XML maintains audit trail data elements within directories that are stored in the Sun and Oracle databases within the Enterprise Computing Center – Memphis (ECC-MEM) (MITS-23, MITS-24). 94x XML is not utilized by users, other than system, server, and database administrators.

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

- A. IRS – The following are a list of systems and applications that are sending data to 94x XML and the data elements that they provide:
- Customer Data Base (CDB)
    - E–File
    - Personal Identification Number (PIN)
    - TIN
    - EIN
  - Electronic Filing System (ELF)
    - Financial Organizations Master File (FOMF) sent by the Electronic Filing System
    - Bank routing numbers
  - National Audit Account (NAP)
    - EIN
    - Name Control –first four characters of last name
  - Third Party Data Store (TPDS)
    - Reporting Agent e–File signatures
- B. Taxpayer:
- Taxpayer information from Forms 940, 941, and 944 are stored in CDB
- C. Employee:
- Usernames
  - Passwords

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

Yes. The data elements are required to carry out 94x XML business purposes, which are to validate 94x XML returns and then process (e.g., extract and reformat) the tax return data for use with other IRS systems.

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

User input to the 94x XML menus is validated against the expected values. If an invalid entry is made, the system displays an “Invalid Choice” message and the user must hit the enter key to be returned to the menu. No processes are executed when invalid responses are entered. Customer Database (CDB), Financial Organizations Master File (FOMF), National Audit Account (NAP), and Third Party Data Store (TPDS) data is checked by system users for the expected format before being loaded into the 94x XML Oracle database tables. 94x XML employs rules to check the valid syntax of information system inputs (e.g., character set, length, numerical range, acceptable values) to verify that inputs match specified definitions for format and content and pre–screens inputs passed to interpreters to prevent the content from being unintentionally interpreted as commands. The CDB, FOMF, NAP, and TPDS data is checked for the correct format before being loaded into the 94x XML Oracle database. If any of these files is found to be in an invalid format, the loading process will abort and the user will be notified of the problem. Tax return data retrieved from EMS (Electronic Management System) is validated according to the methods described in the 94x XML Functional Specification Package document as described above.

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

No, there are no other sources for the data.

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

System administrators located at ECC–MEM access the 94x XML Sun server and subsequently the 94x XML menu using usernames and passwords provided by the MITS–23 and MITS–24 General Support Systems (GSS), which house the 94x XML Unix and Unisys hardware. Users of 94x XML do not directly view taxpayer data. They are responsible for making sure the data goes through and is automatically verified. The user menu does not provide the ability to view actual PII. The menu gives the users the ability to stop applications, verify data if there are any automated error messages, and start the applications back–up.

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

No, data is not retrievable by a personal identifier.

## **Access to the Data**

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

Users of the 94x XML system are IRS System Administrators located at ECC–MEM who are responsible for executing the 94x XML processing on a nightly basis using an administration menu within a Solaris console window. Users must also be added to the Business Master File (BMF) 94x domain group in order to access the menu, execute 94x XML commands, and view log files in the 94x XML directories. In addition, each 94x XML user is granted a non–administrative database account within 94x XML Oracle database. Development team members are granted access to modify 94x XML programming code and must be given access to the 94x XML instance of IBM Rational ClearCase, which is controlled by the Source Control Document Control (SCDC) Branch at the New Carrollton Federal Building. Development team members cannot read, write, edit, or delete data within the 94x XML. An exception to this exists when there is an emergency (“fire call”) with the system and the development team members are given permission from an IRS Director. The SCDC group grants access and sets read/write/edit/delete permissions for each developer/user within ClearCase. 94x XML is administered by IRS MITS employees.

**Role:** System Administrators (SA)

**Permission:** Add and remove users’ permissions within groups within their server. Process permissions granted by user’s manager. Cannot run 94x application.

**Role:** Enterprise Operations (EOPS) Database Administrators (DBA)

**Permission:** Create application administrator (non–DBA) and database schema accounts within Oracle. DBAs have the ability to insert, update, or delete users and objects in the database (e.g., tables or individual records). DBAs do not make changes to the database without a transmittal from the appropriate party. These transmittals can be received from the application project office or Information Technology Infrastructure (ITI), depending on the nature of the change.

**Role:** Computer Systems Analysts (CSA)/Schedulers

**Permission:** Stop systems scripts via the menu on Appendix A, verifies data if there are any error messages, and starts the applications/processes back up at night. Also run weekly clean–up jobs.

*Note: There are no contractors working on this system.*

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

Access to the 94x XML menu, Sun E15k, Unisys mainframe, and Oracle database is controlled by MITS–23 and MITS–24 Network Administrators. System Administrators must be granted access to the BMF 94x user group in order to execute 94x XML processing.

System Administrators performing maintenance on the 94x XML Sun server, Unisys mainframe, or Oracle database must also go through the Online 5081 (OL5081) process in order to be given an account to perform those duties. The OL5081 form lists mandatory rules for users of IRS information systems. When a user has been approved for access to the application by his/her manager, the OL5081 system sends an email to the user, providing an approval notification. The user then logs into the OL5081 system, reads the “Rules of Behavior” section, and provides an “electronic signature”, acknowledging that he/she has read, understands, and agrees to abide by the “Rules of Behavior”.

## **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.**

The following are a list of IRS systems and applications that are sending data to 94x XML:

- Customer Data Base (CDB) – Data is retrieved daily by 94x XML through the 94x tax return data provided by the Electronic Management System (EMS); CDB Type of Data: eFile PIN, TIN, and EIN data; EMS Type of Data: 94x tax return data.
- Electronic Filing System (ELF) – ELF retrieves the Financial Organizations Master File (FOMF) on a monthly basis; Type of Data: Bank routing numbers.
- National Audit Account (NAP) – NAP replaces the contents of the NAP table on a monthly basis with the data retrieved by the retrieve\_NAP.sh process; Type of Data: EIN and Name Control.
- Third Party Data Store (TPDS) – TPDS replaces the contents of TPDS\_EFIN and TPDS\_SW tables on a daily basis with the data from files placed on the 94x XML Sun server by TPDS. This data includes: Reporting Agent e–File signatures, Type of Data: Reporting Agent e–File signatures.

The following are IRS systems/applications that receive data from 94x XML:

- Combined Annual Wage Reporting (CAWR) – is a document matching program that receives data from 94x XML.
- Electronic Federal Tax Payment System (EFTPS) – is an electronic payment system that receives data from 94x XML.
- Electronic Online–Output Network System (EONS) – receives data from 94x XML.
- Electronic Tax Administration Research and Analysis System (ETARAS) – receives data from 94x XML to provide research information about electronic filing.
- Generalized Mainline Framework (GMF) – validates and perfects data that is received from 94x XML before being passed on for master file posting.
- Tax Return Database (TRDB) – receives and stores data from 94x XML on electronically filed tax returns.

94x XML is interconnected to the IRS GSSs and applications identified below:

- Modernization and Information Technology Services (MITS)–1 – IRS Perimeter Security and Network Backbone
- MITS–17 – Enterprise System Management
- MITS–23 – Production, Development, and Test Environment for UNISYS
- MITS–24 – Development and Test Environment for UNIX Consolidated
- MITS–26 – Enterprise Virtual Private Network (VPN)

There are three types of audit logs that are located within Enterprise Computing Center – Memphis (ECC–MEM) MITS–23, MITS–24:

- Oracle database logs, consisting of:
  - Date and time that the event occurred;
  - The unique identifier (e.g., user name, SEID, application name, etc) of the user or application initiating the event;
- 94x XML processing logs on the Sun system, consisting of:
  - Date and time that the event occurred;
  - The unique identifier (e.g., user name, SEID, application name, etc) of the user or application initiating the event;
  - Type of event; and
  - The outcome status (success or failure) of the event.

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

Systems/Applications sending data to 95x XML:

Electronic Management System (EMS)/ Customer Data Base (CDB)<sup>1</sup>

- Certification & Accreditation (C&A) – June 02, 2008
- Privacy Impact Assessment (PIA) – December, 04, 2007

Electronic Filing System (ELF)/ Financial Organizations Master File (FOMF)<sup>2</sup>

- Certification & Accreditation (C&A) – May 26, 2009
- Privacy Impact Assessment (PIA) – April 15, 2009

National Audit Account (NAP)

- Certification & Accreditation (C&A) – May 13, 2009
- Privacy Impact Assessment (PIA) – December 01, 2008

Third Party Data Store (TPDS)

- Certification & Accreditation (C&A) – February 24, 2006

Systems/Applications receiving data from 94x XML:

Combined Annual Wage Reporting (CAWR)

- Certification & Accreditation (C&A) – November 20, 2007
- Privacy Impact Assessment (PIA) – June 14, 2007

Electronic Federal Tax Payment System (EFTPS)

- Certification & Accreditation (C&A) – April 02, 2008
- Privacy Impact Assessment (PIA) – April 02, 2008

Electronic Online–Output Network System (EONS)

- Privacy Impact Assessment (PIA) – April 04, 2002

Electronic Tax Administration Research and Analysis System (ETARAS)

- Certification & Accreditation (C&A) – May 18, 2007
- Privacy Impact Assessment (PIA) – April 13, 2009

### Generalized Mainline Framework (GMF)

- Certification & Accreditation (C&A) – February 18, 2009
- Privacy Impact Assessment (PIA) – October 31, 2008

### Tax Return Database (TRDB)

- Certification & Accreditation (C&A) – May 18, 2007
- Privacy Impact Assessment (PIA) – April 13, 2009

<sup>1</sup>Customer Data Base (CDB) – does not need a Certification & Accreditation (C&A) or a Privacy Impact Assessment (PIA) because it is located within Electronic Management System (EMS)

<sup>2</sup>Financial Organizations Master File (FOMF) – does not need a C&A or a PIA because it is located within the Electronic Filing System (ELF)

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

No. There are no other agencies that provide, receive or share data with 94x XML.

### **Administrative Controls of Data**

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

There are four time frames for eliminating data from 94x XML:

- a) Tax information processed by 94x XML and information generated by the 94x XML processes are not permanently stored on the Sun E15k or Unisys mainframe. 94x XML processing log files are retained for no more than one week from the time the processing takes place.
- b) The following files are not stored longer than one week on the 94x XML system, as they are deleted by the daily/weekly cleanup processes that are initiated by an ECC–MEM System Administrator at midnight six days per week (12:00am, Monday – Saturday) through the 94x XML interfaces:
  - Tax records retrieved from EMS
  - Processing log files (e.g., Build process log, Transfer process log)
  - Tape–backup Archive (TAR) files created by the daily cleanup process (this file contains all of the records which have been moved to ETARAS (Electronic Tax Administration Research and Analysis System) and Unisys as well as all of the rejection/acceptance notices sent back to EMS each day).
- c) 94x XML Unisys output files are stored for 10 Unisys processing cycles before being deleted. After each cycle, the oldest of the 10 stored files (for each file type) is deleted and the new file generated that day is added.
- d) 94x XML Oracle database audit trails and host GSS operating system audit logs are retained for six (6) years, unless otherwise specified by a formal Records Retention Schedule developed in accordance with Internal Revenue Manual (IRM) 15.1, Records Management. These files are controlled by ECC–MEM System Database Administrators.

Database audit data is not required to be local to the database for the period of retention, but is available for historical analysis if needed. Audit data is only readable by personnel authorized by Security Specialists (SecSpec). At the end of the retention period, the audit logs are reviewed to determine if the logs require archival at the Federal Records Center or if they

require destruction. Additional guidance is provided in IRM 1.15. A request for records disposition authority for 94x XML 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 94x XML inputs, system data, outputs and system documentation will be published under IRM 1.15.29 Submissions Processing Campus Records (item number to be determined)

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

No, the system 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, the system 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, the system does not provide the capability to monitor individuals or groups.

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

No, the system does not allow 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. This system does not make any determinations affecting any parties.

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

Not applicable. This is not a web-based application.

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