Date of Approval: 11/03/2017 PIA ID Number: **2800**

A. SYSTEM DESCRIPTION

1. Enter the full name and acronym for the system, project, application and/or database. Cybersecurity Data Warehouse , CSDW

2. Is this a new system? No

2a. If **no**, is there a PIA for this system? Yes

If **yes**, enter the full name, acronym, PIA ID Number and milestone of the most recent PIA.

Cybersecurity Data Warehouse, CSDW, PCLIA #1011

Next, enter the **date** of the most recent PIA. 8/13/2014

Indicate which of the following changes occurred to require this update (check all that apply).

Yes	Addition of PII
No	Conversions
No	Anonymous to Non-Anonymous
No	Significant System Management Changes
No	Significant Merging with Another System
Yes	New Access by IRS employees or Members of the Public
No	Addition of Commercial Data / Sources
Yes	New Interagency Use
Yes	Internal Flow or Collection

Were there other system changes not listed above? No

3. Check the current ELC (Enterprise Life Cycle) Milestones (select all that apply)

No	Vision & Strategy/Milestone 0
No	Project Initiation/Milestone 1
No	Domain Architecture/Milestone 2
No	Preliminary Design/Milestone 3
No	Detailed Design/Milestone 4A
No	System Development/Milestone 4B
Yes	System Deployment/Milestone 5
No	Operations & Maintenance (i.e., system is currently operational)

4. Is this a Federal Information Security Management Act (FISMA) reportable system? Yes

A.1 General Business Purpose

5. What is the general business purpose of this system? Provide a clear, concise description of the system, application or database, the reason for the system, and the benefits to the IRS to use the information, and how the information will be used. The Cyber Security Data Warehouse (CSDW) acts as a centralized data collection and repository point for all Enterprise system logs/syslog's collection. Routers, Servers, Appliances and Applications generate logs which are then forwarded to the Cyber Security Data Warehouse (CSDW). These logs are stored in a central location where Cyber analysts access them and perform analytics using various software security tools. They perform this analysis for the purpose of correlation and identification of fraudulent transactions. The log data can contain personally identifiable information (PII) and sensitive but unclassified (SBU) data such as system/application access, phone numbers, dates of birth, addresses, number of dependents and social security numbers. The business objective of the Cyber Security Fraud Analytics and Monitoring (CFAM) team is to detect and prevent identity theft and frauds. Gathering the log information is also essential so we can refer identified fraud or Identity theft cases to other IRS stakeholders: Return Integrity and Compliance Services (RICS), Criminal Investigation (CI) and Treasury Inspector General for Tax Administration (TIGTA).

B. PII DETAIL

6. Does the system use, collect, receive, display, store, maintain, or disseminate IR Code 6103 taxpayer information; or any type of Sensitive but Unclassified (SBU) or Personally Identifiable Information (PII)? Yes

6a. If **yes**, does the system use, collect, receive, display, store, maintain, or disseminate Social Security Numbers (SSN s) or tax identification numbers (i.e. last 4 digits, etc.)? Yes

If **yes**, check who the SSN (or tax identification number) is collected on.

Yes	On Primary	Yes	On Spouse	Yes	On Dependent

If **yes**, check all types SSN s (or tax identification numbers) that apply to this system:

Yes	Social Security Number (SSN)
Yes	Employer Identification Number (EIN)
Yes	Individual Taxpayer Identification Number (ITIN)
No	Taxpayer Identification Number for Pending U.S. Adoptions (ATIN)
Yes	Practitioner Tax Identification Number (PTIN)

Describe the planned mitigation strategy and forecasted implementation date to mitigate or eliminate the use of SSN s (or tax identification numbers). Due to the nature of the Fraud analytics conducted against the data, mitigation or elimination of SSN's is not possible. The Office of Management and Budget memorandum M-07-12 requires that federal agencies develop a mitigation or elimination strategy for systems that use SSNs, which the Service continues to develop strategies to meet. An exception to that requirement is when the SSN is uniquely needed to identify a user's record. The Cybersecurity Fraud Analytics and Monitoring program requires the use of SSN's because no other identifier can be used to uniquely identify a taxpayer at this time. SSNs are permissible from Internal Revenue Code (IRC) 6109, which requires individual taxpayers to include their SSNs on their income tax returns.

6b. Does this system contain other (non-SSN) PII that it uses, collects, receives, displays, stores, maintains, or disseminates? (i.e. Names, addresses, etc.) Yes

If **yes**, specify the information.

Selected	PII Element	On Primary	On Spouse	On Dependent
Yes	Name	Yes	Yes	Yes
Yes	Mailing address	No	No	No
Yes	Phone Numbers	No	No	No
Yes	E-mail Address	No	No	No
Yes	Date of Birth	Yes	Yes	Yes
No	Place of Birth	No	No	No
Yes	SEID	No	No	No
No	Mother's Maiden Name	No	No	No
Yes	Protection Personal Identification Numbers (IP PIN)	No	No	No
Yes	Internet Protocol Address (IP Address)	No	No	No
No	Criminal History	No	No	No
No	Medical Information	No	No	No
No	Certificate or License Numbers	No	No	No
No	Vehicle Identifiers	No	No	No
Yes	Passport Number	No	No	No
Yes	Alien (A-) Number	No	No	No
Yes	Financial Account Numbers	No	No	No
No	Photographic Identifiers	No	No	No
No	Biometric Identifiers	No	No	No
Yes	Employment (HR) Information	No	No	No
Yes	Tax Account Information	Yes	Yes	Yes

6c. Does this system contain SBU information that is not PII, it uses, collects, receives, displays, stores, maintains, or disseminates? Yes

If **yes**, select the types of SBU

Selected	SBU Name	SBU Description
Yes	Agency Sensitive Information	Information which if improperly used or disclosed could adversely affect the ability of the agency to accomplish its mission
No	Procurement sensitive data	Contract proposals, bids, etc.
Yes	Official Use Only (OUO) or Limited Official Use (LOU)	Information designated as OUO or LOU is information that: is exempt under one of the statutory Freedom of Information Act exemptions; is prohibited by other laws or regulations; would significantly impede the agency in carrying out a responsibility or function; or would constitute an unwarranted invasion of privacy.
No	Proprietary data	Business information that does not belong to the IRS
No	Protected Information	Information which if modified, destroyed or disclosed in an unauthorized manner could cause: loss of life, loss of property or funds by unlawful means, violation of personal privacy or civil rights, gaining of an unfair procurement advantage by contractors bidding on government contracts, or disclosure of proprietary information entrusted to the Government
No	Physical Security Information	Security information containing details of serious weaknesses and vulnerabilities associated with specific systems and facilities
Yes	Criminal Investigation Information	Information concerning IRS criminal investigations or the agents conducting the investigations.

6d. Are there other types of SBU/PII used in the system? Yes

If **yes**, describe the other types of SBU/PII that are applicable to this system. Sensitive But Unclassified briefings. Logs containing information related to Network Internet Protocol address (IP address). Server names. A Uniform Resource Locator (URL), also referred to as a web address. Domain Name Servers (DNS).

6e. Cite the authority for collecting SBU/PII (including SSN if relevant)

Yes PII for federal tax administration is generally Internal Revenue Code Sections 6001, 6011, 6012e(a)
Yes SSN for tax returns and return information is Internal Revenue Code Section 6109
No SSN for personnel administration (IRS Employees) is 5 USC & Executive Order 9397
Yes PII for personnel administration is 5 USC
No PII about individuals for Bank Secrecy Act compliance 31 USC
Yes Information by CI for certain money laundering cases may be 18 USC

6f. Has the authority been verified with the system owner? Yes

B.1 BUSINESS NEEDS AND ACCURACY

7. Explain the detailed business needs and uses for the SBU/PII, and how the SBU/PII is limited only to that which is relevant and necessary to meet the mission requirements of the system. If SSNs (or tax identification numbers) are used, explicitly explain why use of SSNs meets this criteria. Be specific. The business objective of Cybersecurity Fraud Analytics and Monitoring (CFAM) is to detect and prevent identity theft and frauds. To achieve this, CFAM relies on the monitoring and analyzing of log transactions with IRS's online services. These transactions records contain PII/SBU including: SSNs/TINs, Internet Protocol (IP) addresses, email addresses, phone numbers, refund amounts, home addresses, etc. They are the core data fields for CFAM's analytic and modeling effort to detect suspicious access patterns and identify potential Identity Theft (IDT)/fraud perpetrators and victims. SSN/TIN: the main item CFAM users to identify a taxpayer and extract online transactions related to his/her account. CFAM also receives SSN/TIN information from RICS/CI/TIGTA when they request CFAM's collaboration in various investigations. Additionally, CFAM also uses SSN/TIN to request further information from IRS organizations/databases, such the Return Review Program (RRP) or Compliance Data Warehouse (CDW). Finally, CFAM also uses SSN/TIN to extract family clusters and detect abnormal filing patterns. Internet Protocol Address (IP addresses)/Email addresses/Phone numbers: CFAM uses these items to detect and identify suspicious online accesses and activities. For example, suspicious actors tend to use one IP address/email account/phone number to access many taxpayers' accounts. Refund amounts: CFAM has found that suspicious actors tend to have abnormally distributed refund amounts, and are good indicators of large-scale refund fraud. Home addresses: CFAM uses such info to compare with the geolocation of the IP addresses and users any discrepancies for potential leads for IDT/fraud detection. For these activities, the PII/SBU items listed above are relevant and necessary to achieve CFAM's mission.

8. How is the SBU/PII verified for accuracy, timeliness, and completeness? Explain how steps are taken to ensure that all information maintained by the system that is used by IRS to make any adverse determination about an individual's rights, benefits, and/or privileges is maintained with such accuracy, relevance, timeliness, and completeness as is reasonably necessary to assure fairness to the individual in the determination. The data that Cybersecurity Data Warehouse receives is from internal IRS systems which are deemed reliable and the data is validated for accuracy by the system sending the data as described in that system's PCLIA. Any determinations made are validated during the analysis process and analysts can work directly with system owners to verify integrity. The stored data Cybersecurity Fraud Analytics and Monitoring (CFAM) uses, is presented to the analyst in a read-only format, therefore CFAM's analysis does not affect the integrity of the raw data.

C. PRIVACY ACT AND SYSTEM OF RECORDS

9. Are 10 or more records containing SBU/PII maintained, stored, and/or transmitted by or through this system? Yes

9a. If **yes**, are records in the system retrieved by any personal identifier (e.g., name, SSN, Photograph, IP Address) for an individual? Yes

If **yes**, is there a System of Records Notice(s) or SORNs that addresses the PII records in this system? Yes

If **yes**, enter the SORN number(s) and the complete the name of the SORN.

SORNS Number	SORNS Name
Treas/IRS 34.037	Audit Trail and Security Records System

If **yes**, does the System of Records Notice(s) (SORN) published in the Federal Register adequately describe the records as required by the Privacy Act? Yes

D. RESPONSIBLE PARTIES

10. Identify the individuals for the following system roles. ## Official Use Only

E. INCOMING PII INTERFACES

11. Does the system receive SBU/PII from other system or agencies? Yes

11a. If **yes**, does the system receive SBU/PII from IRS files and databases? Yes

If **yes**, enter the files and databases.

System Name	Current PIA?	PIA Approval Date	SA & A?	Authorization Date
eAuthentication	Yes	12/30/2015	Yes	10/13/2015
Transcript Delivery System	Yes	11/03/2015	Yes	03/21/2017
Return Review Program	Yes	01/23/2015	Yes	06/23/2017
Integrated Customer Communications Environment	No		No	06/23/2017

11b. Does the system receive SBU/PII from other federal agency or agencies? No

11c. Does the system receive SBU/PII from State or local agency(s)? No

11d. Does the system receive SBU/PII from other sources? Yes

If **yes**, identify the source that sends the SBU/PII, how the SBU/PII is transmitted and if there is an Inter-Agency Agreement (ISA) /Memorandum of Understanding (MOU).

Organization Name	Transmission method	ISA/MOU
Commercial Off the Shelf (COTS) security and security-related applications	Data Collection	Yes

11e. Does the system receive SBU/PII from **Taxpayer** forms? Yes

If **yes**, identify the forms

Form Number	Form Name
1040	1040
1040A	1040A
1040EZ	1040EZ

11f. Does the system receive SBU/PII from **Employee** forms (such as the I-9)? No

F. PII SENT TO EXTERNAL ORGANIZATIONS

12. Does this system disseminate SBU/PII? Yes

12a. Does this system disseminate SBU/PII to other IRS Systems? No

12b. Does this system disseminate SBU/PII to other Federal agencies? No

12c. Does this system disseminate SBU/PII to State and local agencies? No

12d. Does this system disseminate SBU/PII to IRS or Treasury contractors? No

12e. Does this system disseminate SBU/PII to other Sources? Yes

If **yes**, identify the other source(s) that receive SBU/PII from this system, and if there is an Inter-Agency Agreement (ISA) /Memorandum of Understanding (MOU).

Organization Name	Transmission method	ISA/MOU
Privacy, Governmental Liaison and Disclosure's	encrypted email	No
Criminal Investigation	encrypted email	No
Return Integrity & Compliance Services	encrypted email	No
Treasury Inspector General for Tax Administration	encrypted email	No

Identify the authority and for what purpose? The authority was given by the ACIO of Cybersecurity and the Director of Security Operations. The PII data is shared for the purpose of supporting criminal investigations, identifying potential tax refund fraud, and facilitating treatment for the taxpayers that were victimized. TIGTA - Treasury Inspector General for Tax Administration CI - Criminal Investigation RICS - Return Integrity & Compliance Services PGLD - Privacy, Governmental Liaison and Disclosure's.

G. PRIVACY SENSITIVE TECHNOLOGY

13. Does this system use social media channels? No

14. Does this system use privacy-sensitive technologies such as mobile, global position system (GPS), biometrics, RFID, etc.? No

15. Does the system use cloud computing? No

16. Does this system/application interact with the public? No

H. INDIVIDUAL NOTICE AND CONSENT

17. Was/is notice provided to the individual prior to collection of information? Yes

17a. If **yes**, how is notice provided? Was the individual notified about the authority to collect the information, whether such is mandatory or voluntary, the purpose for which the information will be used, with whom the information may be shared, and the effects, if any, if they decide not to provide any of the requested information? The IRS notifies all individuals who file tax returns of such collection via the Privacy Act Notice in tax return instructions. Our legal right to ask for information is Internal Revenue Code sections 6001, 6011, and 6012(a), and their regulations. They say that taxpayers must file a return or statement with us for any tax they are liable for. Their response is mandatory under these sections".

18. Do individuals have the opportunity to decline from providing information and/or from consenting to particular uses of the information? No

18b. If no, why not? CyberSecurity Data Warehouse does not directly provide individuals the opportunity to decline from providing information and/or from consenting to particular uses of the information. Notice, consent and due process are provided in the tax forms instructions filed by the taxpayer, and pursuant to 5 USC. Or if gathered from tax form: The IRS has the legal right to ask for information per Internal Revenue Code sections 6001, 6011, and 6012(a), and their regulations. The regulations state that "taxpayers must file a return or statement with IRS for any tax they are liable for". Their response is mandatory under these sections.

19. How does the system or business process ensure due process regarding information access, correction and redress? The CyberSecurity Data Warehouse process and procedures are dictated by the Internal Revenue Manual guidelines. IRS policy allows individuals whose data is in the system the opportunity to clarify or dispute negative determinations through the appeals process.

I. INFORMATION PROTECTION

20. Identify the owner and operator of the system (could be IRS owned and Operated; IRS owned, contractor operated; contractor owned and operated). IRS Owned and Operated

21. The following people have access to the system with the specified rights:

IRS Employees? Yes

IRS Employees?	Yes/No	Access Level (Read Only/Read Write/ Administrator)
Users	Yes	Read-Only
Managers	Yes	Read-Only
Sys. Administrators	Yes	Read and Write
Developers	Yes	Read-Only

Contractor Employees? Yes

Contractor Employees?	Yes/No	Access Level	Background Invest. Level
Contractor Users	Yes	Read-Only	Moderate
Contractor Managers	Yes	Read-Only	Moderate
Contractor Sys. Admin.	No		
Contractor Developers	Yes	Read-Only	Moderate

21a. How is access to SBU/PII determined and by whom? The Cybersecurity Data Warehouse system utilizes the IRS On–Line application OL-5081 application to document approvals for access. Data access is granted on a need-to-know basis. A potential user must submit a request for access via IRS OL5081 to their local management for approval consideration. Users are not permitted access without a signed Form 5081 from an authorized management official. Specific permissions (Read, Write, Modify, Delete, and/or Print) are defined on the OL5081 form and set (activated) by the System Administrator prior to the user being allowed access. User privileges and user roles determine the types of data that each user has access to. Management monitors system access and removes permissions when individuals no longer require access. Users are assigned to specific modules of the application and specific roles within the modules and accounts follow the principle of least privilege which provide them the least amount of access to PII data that is required to perform their business function after receiving appropriate approval.

21b. If computer matching occurs, can the business owner certify that it meets requirements of IRM 11.3.39 Disclosure of Official Information, Computer Matching & Privacy Protection Act?

Not Applicable

I.1 RECORDS RETENTION SCHEDULE

22. Are these records covered under the General Records Schedule (GRS), or have a National Archives and Records Administration (NARA) archivist approved a Record Control Schedule (RCS) for the retention and destruction of official agency records stored in this system? Yes

22a. If **yes**, how long are the records required to be held under the corresponding RCS and how are they disposed of? In your response, please include the complete IRM number 1.15.XX and specific item number and title. CSDW extracts data from sources. Potential data retention guidance: IRM 1.15.17. CSIRC analysts require retention of all CSDW sources for 2 years for analytical purposes. TIGTA and FISMA require retention of critical system logs for up to seven years and standard security logs for up to four years. All records housed in the CSDW system will be erased or purged from the system in accordance with approved retention periods. It is the official repository for data and documents and has National Archives approval to affect data disposition. Any records generated and maintained by the system will be managed according to requirements under IRM 1.15.1 and 1.15.6, and will be destroyed using IRS General Records Schedule (GRS) 3.2 for Information System Security Records, Items 010, 030 and as coordinated with the IRS Records and Information Management (RIM) Program and IRS Records Officer.

I.2 SA&A OR ECM-R

23. Has the system been through SA&A (Security Assessment and Authorization) or ECM-R (Enterprise Continuous Monitoring Reauthorization)? No

23c. If **no**, is the system secured in accordance with all applicable federal, treasury, and IRS security policy, procedures, and requirements? Do not know

23.1 Describe in detail the system s audit trail. Currently audit is captured via system logging. Procurements are being made to add additional audit capabilities.

J. PRIVACY TESTING

24. Does the system require a System Test Plan? No

24c. If **no**, please explain why. Continuous Monitoring (eCM)(now called Annual Security Control Assessment (ASCA)) occurs annually to ensure that controls remain in place to properly safeguard PII.

K. SBU Data Use

25. Does this system use, or plan to use SBU Data in Testing? Yes

25a. If **yes,** was permission granted per the requirements of Form 14664, SBU Data Use Questionnaire or Form 14665, SBU Data Use Request? No

If **no**, explain why not. Build out of dev environment is currently under way for future upgrade of Hadoop environment. Forms will need to be completed prior to using SBU for testing.

L. NUMBER AND CATEGORY OF PII RECORDS

26. Identify the number of individual records in the system for each category:

26a. IRS Employees: More than 100,000
26b. Contractors: Under 5,000
26c. Members of the Public: More than 1,000,000
26d. Other: Yes

If **other**, identify the category of records and the number of corresponding records (to the nearest 10,000). FATCA - Foreign Account Tax Compliance Act and related services listed below. At this time the number of records is undetermined as analytics is not currently being performed on this data. Fatca AccountSummaryService Fatca BranchService Fatca FIService Fatca FITransferService Fatca MemberDetailsService Fatca MemberService Fatca MessageService Fatca PAIService Fatca PersonService Fatca SEBranchService Fatca SponsoredEntityService Fatca SubmitRegistrationService Fatca UserAcctService

M. CIVIL LIBERTIES

27. Does the system maintain any information describing how any individual exercises their rights guaranteed by the First Amendment? Yes

27a. If **yes**, explain the First Amendment information being collected and how it is used. The system collects information supplied via Tax forms to validate Taxpayers are who they say they are when they are requesting their Taxpayer documents. These factors are used as a validation process.

27b. If **yes**, please check all of the following exceptions (any one of which allows the maintenance of such information) that apply:

The individual about whom the information was collected or maintained expressly authorizes its collection/maintenance (as noted in Q17). Yes
The information maintained is pertinent to and within the scope of an authorized law enforcement activity. (As noted in Q 7) Yes

There is a statute that expressly authorizes its collection. (Identified in Q6) No

27c. If **yes**, will the First Amendment information be used as the basis to make any adverse determination about an individual's rights, benefits, and/or privileges? No

28. Is the system information used to conduct data-mining as defined in the *Implementing the 9/11 Commission Recommendations Act of 2007, Public Law 110-53, Section 804?* No

29. Will this system have the capability to identify, locate, and monitor individuals or groups of people? No

N. ACCOUNTING OF DISCLOSURES

30. Does the system include or require disclosure of tax or employee information to anyone other than IRS employees in the performance of their duties, or to the person to whom the information pertains or to a 3rd party pursuant to a Power of Attorney, tax or Privacy Act consent? No

End of Report