Date of Approval: **July 12, 2022**

PIA ID Number: **7125**

SYSTEM DESCRIPTION

Enter the full name and acronym for the system, project, application and/or database.

Enhanced Paper Processing, EPP

Is this a new system?

No

Is there a PCLIA for this system?

Yes

What is the full name, acronym, and milestone of the most recent PCLIA?

Enhanced Paper Processing

What is the approval date of the most recent PCLIA?

9/16/2021

Changes that occurred to require this update:

Significant System Management Changes

Significant Merging with Another System

Were there other system changes not listed above?

No

What governance board or Executive Steering Committee (ESC) does this system report to? Full name and acronym.

This is an external vendor system run by IBM under IRS contract 47QTCK18D0014. It does not fall under internal IRS governance board procedures.

Current ELC (Enterprise Life Cycle) Milestones:

Operations & Maintenance (i.e., system is currently operational)

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

No

GENERAL BUSINESS PURPOSE

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 purpose of the Enhanced Paper Processing (EPP) is to convert paper tax returns into an electronic format. The process starts with document preparation (preparing the paper returns and schedules for scanning which includes visual inspection of documents for tears, dog-ears, hole punches that may interfere, pulling staples and other fasteners, repairing documents, reordering as necessary). Prepared documents are then scanned using high-speed scanning/Softrac Image Enhance/Output, which in turn are ingested into Appian and Hyperscience components of the extended solution hosted on AWS GovCloud via a watch folder over a secure VPN connection between the IRS enclave on Allegany Ballistics Laboratory (ABL) network and AWS GovCloud. There is no direct connection between AWS GovCloud and the IRS. In the GovCloud enclave the Hyperscience component performs forms recognition, Optical Character Recognition (OCR), and field zoning/metadata extraction, on the now-electronic tax returns. It captures certain batch level data from batch transmittal sheets and pairs that information with file naming schema from the scanning process to establish Hyperscience batch naming schema. Each return becomes its own batch and the established rule sets (project requirements) dictate the targeted metadata fields for extraction to arrive at a Modernized e-File (MeF) compliant XML file for acceptance into downstream IRS return approval/determination processes. While the extraction of the data from the documents and basic field/data validation are handled through the Hyperscience component of the solution, the execution of business rules to validate whether all required info is available based on schema and creation of the target MeF compliant XML file is handled by the Appian component of the solution in the AWS GovCloud. During OCR extraction process, based on low extraction confidence score and error scenarios, some documents are routed to IBM reviewers in the loop staff to address them through the Hyperscience web application. Similarly, documents that require human intervention/supervision during data validation for MeF compliant XML, are routed to an IRS Specialist to review and correct through the Appian web application. IRS Specialists will access the Appian browser-based application through the Virtual Private Network (VPN) connection used to access the server at the Rocket Center, but no data will be pulled back or stored on the IRS laptop. A zero-footprint browser, Appian, will be configured to prevent any downloading or uploading of data from the IRS. All work is performed on the IBM ABL remotely by the IRS Specialists. The AWS GovCloud enclave does not have any direct connectivity to the IRS network or its users. All extracted, validated, and generated MeF compliant XML returns are shared with the TASPR (Data Transcription Tool that massages the data for ingestion)/MeF through network file share over the VPN connection between IBM ABL and AWS GovCloud via the IBM ABL VPN connection to the IRS. These same files and any native files (raw scanned images, batch info, log files) remain on the IRS

enclave until approved for deletion (release) by IRS via determinations from reconciliation reports. Upon approval, the native paper files are prepared for return to designated IRS centers via IRS-provided transportation service providers and the electronic files are purged (released) from IBM network to free up space for future capacity for the project. The AWS GovCloud enclave does not have any direct connectivity to IRS network.

PII DETAILS

Does the system use, collect, receive, display, store, maintain, or disseminate IR Code 6103 taxpayer information: or any other type of Sensitive but Unclassified (SBU) information or PII such as information about IRS employees or outside stakeholders?

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

What types of tax identification numbers (TIN) apply to this system?

Social Security Number (SSN)

List the approved Treasury uses of the SSN:

Another compelling reason for collecting the SSN

Explain why the authorized use(s) above support the new or continued use of SSNs (or tax identification numbers).

The EPP system processes IRS 1040 tax forms which contain Taxpayer SSNs. All PII information on a tax submission is collected and processed by EPP. Modification of tax submissions to omit any data would modify the original source of the tax record.

Describe the planned mitigation strategy and forecasted implementation date to mitigate or eliminate the use of SSN's (or tax identification numbers).

The EPP system requires the use of SSN's because no other identifier can be used to uniquely identify a taxpayer. SSNs are permissible from Internal Revenue Code (IRC) 6109, which requires individual taxpayers to include their SSNs on their income tax returns.

Employer Identification Number

Other Taxpayer Identification Number

Does this system use, collect, receive, display, store, maintain or disseminate other (non-SSN) PII (i.e., names, addresses, etc.)?

Yes

Specify the PII Elements:

Name
Mailing address
Phone Numbers
Date of Birth
Place of Birth
Standard Employee Identifier (SEID)
Mother's Maiden Name
Protection Personal Identification Numbers (IP PIN)
Certificate or License Numbers
Financial Account Numbers
Tax Account Information
Centralized Authorization File (CAF)

Does this system use, collect, receive, display, store, maintain, or disseminate SBU information that is not PII?

Yes

Specify the types of SBU from the SBU Types List:

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.

Are there other types of SBU/PII used in the system?

Yes

Describe the other types of SBU/PII that are applicable to this system.

Taxpayer Data (Federal Tax Information [FTI])

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

PII for federal tax administration is generally Internal Revenue Code Sections 6001, 6011, & 6012e(a)

SSN for tax returns and return information is Internal Revenue Code Section 6109

Has the authority been verified with the system owner?

Yes

BUSINESS NEEDS AND ACCURACY

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 taxpayer's information is required for each tax return. Since the purpose of EPP is to scan in and digitize paper taxpayer 1040 forms, all taxpayer information will be included, to include SSN.

How is the SBU/PII verified for accuracy, timeliness, and completion?

After the paper forms are scanned in through the Enhanced Paper Processing project, they are prepared as electronic returns and go through document perfection to be transitioned to the MeF-compliant format so that they can be electronically submitted. Each return prepared and submitted to MeF for e-filing must adhere to the schemas and business rules. If a single data element fails the schema integrity check or business rule failure, the tax return is rejected. Electronic Tax Administration (ETA) supplies the business rules for each return type. MeF enforces the rules against the tax returns using a business rules engine. Business rules enforce relationships between data and forms. When MeF validates returns against the business rules, and if it encounters a discrepancy, the tax return is rejected. New schemas and business rules are issued for each new tax year. Any non-current returns (prior tax years) prepared and submitted must use that year's schema and business rule versions or the returns will fail the schema and business rule validation checks.

PRIVACY ACT AND SYSTEM OF RECORDS

The Privacy Act requires Federal agencies that maintain a system of records to publish systems of records notices (SORNs) in the Federal Register for records from which information is retrieved by any personal identifier for an individual who is a US citizen, or an alien lawfully admitted for permanent residence. The Privacy Act also provides for criminal penalties for intentional noncompliance.

Does your application or this PCLIA system pertain to a group of any record from which information is retrieved by any personal identifier for an individual who is a US citizen, or an alien lawfully admitted for permanent residence? An identifier may be a symbol, voiceprint, SEID, or other personal identifier that is used to retrieve information.

Yes

Identify the Privacy Act SORN(s) that cover these records.

IRS 22.062   Electronic Filing Records

IRS 34.037   Audit Trail and Security Records

IRS 24.030   Customer Account Data Engine Individual Master File

IRS 24.046   Customer Account Data Engine Business Master File

IRS 00.001   Correspondence Files and Correspondence Control Files

RESPONSIBLE PARTIES

Identify the individuals for the following system roles:

Official Use Only

INCOMING PII INTERFACES

Does the system receive SBU/PII from other systems or agencies?

Yes

Does the system receive SBU/PII from IRS files and databases?

No

Does the system receive SBU/PII from other federal agency or agencies?

No

Does the system receive SBU/PII from State or local agency (-ies)?

No

Does the system receive SBU/PII from other sources?

No

Does the system receive SBU/PII from Taxpayer forms?

Yes

Please identify the form number and name:

Form Number: 1040
Form Name: US Individual Income Tax Return

Does the system receive SBU/PII from Employee forms (e.g., the I-9)?

No

DISSEMINATION OF PII

Does this system disseminate SBU/PII?

Yes

Does this system disseminate SBU/PII to other IRS Systems?

Yes

Identify the full name and acronym of the IRS system(s) that receive SBU/PII from this system.

System Name: Modernized eFile
Current PCLIA: Yes
Approval Date: 2/8/2022
SA&A: Yes
ATO/IATO Date: 6/14/2022

Identify the authority.

E-Government Act of 2002, 5 USC 301 or 26 USC 6103(b)(1) and IRC 6103(n

For what purpose?

Electronically submit tax returns

Does this system disseminate SBU/PII to other Federal agencies?

No

Does this system disseminate SBU/PII to State and local agencies?

No

Does this system disseminate SBU/PII to IRS or Treasury contractors?

No

Does this system disseminate SBU/PII to other Sources?

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: AWS GovCloud (West)
Transmission Method: secure VPN
ISA/MOU: Yes

Identify the authority.

E-Government Act of 2002, 5 USC 301 or 26 USC 6103(b)(1) and IRC 6103(n)

Identify the Routine Use in the applicable SORN (or Privacy Act exception).

No change to previous PCLIA 22.062 Electronic Filing Records 34.037 Audit Trail and Security Records System 24.030 Individual Master File 24.046 Business Master File

For what purpose?

Leverage AWS GovCloud resources for form recognition, OCR, and field zoning/metadata extraction on the now-electronic tax forms.

PRIVACY SENSITIVE TECHNOLOGY

Does this system use social media channels?

No

Does this system use privacy-sensitive technologies such as mobile, global position system (GPS), biometrics, RFID, etc.?

No

Does the system use cloud computing?

Yes

Is the cloud service provider (CSP) Federal Risk and Authorization Management Program (FedRAMP) certified?

Yes

Date Certified.

6/21/2016

Please identify the ownership of the CSP data.

Third Party

Does the CSP allow auditing?

Yes

Who audits the CSP Data?

3rd Party

What is the background check level required for CSP?

Moderate

Is there a breach/incident plan on file?

Yes

Privacy laws (including access and ownership) can differ in other countries. This cloud will be Continental US (CONUS) only for:

Storage
Transmission

Does this system/application interact with the public?

No

INDIVIDUAL NOTICE AND CONSENT

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

Yes

How is notice provided? Was the individual notified about the authority to collect the information, whether disclosure is mandatory or voluntary, the purpose for which the information will be used, with whom the information may be shared, and the effects on the individual, if any, if they decide not to provide all or any of the requested information?

Form 1040 instructions to the taxpayer.

Do individuals have the opportunity to decline from providing information and/or from consenting to particular uses of the information?

No

Why not?

Information required to electronically file tax returns. Authority: Internal Revenue Code sections 6001, 6011, and 6012(a), and their regulations & Privacy Requirements (PVR) #15-Consent and #18- Individual Rights.

How does the system or business process ensure 'due process' regarding information access, correction, and redress?

Due process is addressed by other IRS business departments that directly interact with taxpayers. Due Process is provided pursuant to Title 26 USC.

INFORMATION PROTECTION

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

Contractor Owned and Operated

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

IRS Employees

Users: Read Write

IRS Contractor Employees

Contractor Users: Read Write

Contractor System Administrators: Administrator

Contractor Developers: Administrator

How is access to SBU/PII determined and by whom?

IRS Employees - access to EPP is restricted to VPN access administered by IBM. W&I Electronic Filing Services (EFS) provides IBM with a list of authorized users for the EPP project and IBM controls access via user account in active directory. Only those IRS users with the appropriate credentials are able to login to the VPN to perform data analysis and

modification. Contractors - access o EPP is restricted to those contractors and subcontractors who have been cleared by the IRS and are associated with the EPP contract. Users are controlled by role-based access via user accounts in active directory, administered by IBM. New accounts requests or modifications are managed via ServiceNow tickets. The IBM Program Manager for EPP approves the need for the appropriate accounts.

RECORDS RETENTION SCHEDULE

Are these records covered under a General Records Schedule (GRS, IRS Document 12829), or has the National Archives and Records Administration (NARA) approved a Records Control Schedule (RCS, IRS Document 12990) for the retention and destruction of official agency records stored in this system?

Yes

How long are the records required to be held under the corresponding GRS or RCS, and how are they disposed of? In your response, please provide the GRS or RCS chapter number, the specific item number, and records series title.

RCS 19 Item 81-Modernized e-File (MeF) System. -(A) Inputs: The Modernized e-File System (MeF) end users register for IRS e- Services program and complete an application for IRS e-file. After an applicant passes the suitability check and the IRS completes processing the application, the IRS notifies the applicant of acceptance to participate in IRS e-file. Transmitters and Software Developers must complete testing before final acceptance. The IRS assigns Electronic Identification Filing Numbers (EFINs) to all Providers and assigns Electronic Identification Transmission Numbers (ETINs) to Transmitters, Software Providers, and Online Providers. The following Form types are e-filed through the MeF platform as inputs: Corporations: Forms 1120, 1120-F, and 1120S. Exempt Organizations: Forms 1120-POL, 990, 990-EZ, 990-N (e-Postcard), and 990- PF. Partnerships: Forms 1065 and 1065-B. Excise Tax: Forms 2290, 720 and 8849 (Schedules 3, 5, 6 only) Extensions: Forms 7004 and 8868. MeF also supports a Fed/State program that consists of a single point of submission and retrieval for all registered transmitters and State agencies. This Fed/State initiative includes Corporations, Partnerships and Exempt Organizations. Additional MeF information can be found in Publications 4163 and 4164. AUTHORIZED DISPOSITION Delete/Destroy any cached transmission files 90 days after successful entry and verification into the MeF Repository. (B) System Data: System data in the Modernized e-File System (MeF) consists of the contents of all transmissions files from taxpayers. AUTHORIZED DISPOSITION Cut off at the end of each processing year. Delete/Destroy 75 years after cutoff. (C) Outputs: The Modernized e-File System (MeF) outputs support eFile and Modernized eFile Programs supported by the System (including Business Master File, Business Master File Returns Processing, End-User (public) documentation hosted on the IRS Internet web site. Documentation includes Internal Revenue Service Publications 3112, 3005, 1345, 1346, 1436, 1437, 1438, 1438-A, 1474, 3823, 4162, 4163, 4164, 4205, 4505, 4594, Revenue Procedure 2007-40 e-Glossary, Quick Alerts, and e-News. AUTHORIZED DISPOSITION Delete/Destroy when no longer needed for operational purposes. Note: MeF

outputs are retained in the recipient electronic publications system as the official records. (D) System Documentation: System Documentation for the Modernized e-File System (MeF) consists of codebooks, records layout, User Guide, and other related materials. AUTHORIZED DISPOSITION Delete/Destroy when superseded or 5 years after the system is terminated, whichever is sooner. Reporting Structure for Return Reconciliation Example Block of Returns (4 returns): IRS-EPP-09-07072021-0005-1052021-001-001-023-004-001-001 IRS-EPP-09-07072021-0005-1052021-001-001-023-004-001-002 IRS-EPP-09-07072021-0005-1052021-001-001-023-004-001-003 IRS-EPP-09-07072021-0005-1052021-001-001-023-004-001-004 Definition of String: IRS-EPP = Project 09 = FLC Service Center 07072021 = Date Shipped 0005 = Batch Number 1052021 = Receive Date 001 = Box Number N 001 = Total Boxes M 023 = Block Number 004 = Returns in Block/Softtrac Batch 001 = Datacap Batch ID (for entire Block/Softtrac Batch) 001-004 = Return Number from Original Datacap Batch (Child batch created to route returns individually to export or Supervisor review) o Box Return Process: ï‚§ Datacap, in coordination with metrics fed by TASPR/MeF, will produce "Completion" and "Return Readiness" Reports. We refer to this as the "Reconciliation Report." ï‚§ IBM Records Management team will use said reports to determine the feasibility and logistics for upcoming return shipments to the IRS Centers In order for a Box to be returned, all returns for the Box must have been successfully loaded into MeF ("Accepted" or "Manual") and any "Bad Returns" have been flagged for return within the parent box, Datacap Bad Return Process closed as Complete, and the status in the Reconciliation Return updated to "Bad Return" from "blank." A copy of the Reconciliation Report is also included within the box for visual inspection, with any "Bad Returns" highlighted. In order for an IRS Batch (9382)/Shipment to be returned, all returns for the Boxes that make up this batch must have been successfully loaded into MeF and any "Bad Returns" have been flagged for return, all appropriate comms have been completed to IRS POCs, and the Box statuses for all associated Boxes in the IRS Batch/Shipment are verified to contain returns that are in either "Bad Return," "Accepted," or "Manual" for its TASPR status within the IBM Reconciliation Report. ï‚§ IBM, in coordination with IRS Center POCs will schedule "Shipment Pickups" according to shipments that have met the above criteria. Unless otherwise directed, the return shipments will be shipped back to IRS Centers as they were originally shipped to IBM. I.e., Shipment 0001 will not be returned until all returns from all boxes contained within that original shipment are complete and ready for return. This represents final chain-of-custody handoff. IRS Third-party transportation vendors may be used, as well as standard couriers such as UPS and FEDEX. o Report Requirements ï‚§ Database must contain all variables of the aforementioned string ï‚§ Database must also contain the current status of any and all returns that have been ingested by Datacap ï‚§ Queries need to be available to be ran at the following levels: 09 = FLC Service Center 07072021 = Date Shipped 0005 = Batch Number 001 = Box Number N 023 = Block Number o Sample Report for Block 023 from above: ï‚§ Query was executed for the above criteria: o IRS-EPP-09-07072021-0005-1052021-001-001-023-004-001-001 "Accepted" o IRS-EPP-09-07072021-0005-1052021-001-001-023-004-001-002 "Accepted" o IRS-EPP-09-07072021-0005-1052021-001-001-023-004-001-003 "Manual" o IRS-EPP-09-07072021-0005-1052021-001-001-023-004-001-004 "Bad Return"

SA&A OR ASCA

Has the system been through SA&A (Security Assessment and Authorization) or ASCA (Annual Security Control Assessment)?

No

Is the system secured in accordance with all applicable federal, treasury, and IRS security policy, procedures, and requirements?

Yes

Describe the system's audit trail.

The system has been assessed under the Cyber Contractor Site Assessment process against the IRS Publication 4812 for Contractor Security Controls. The system audit trail is managed by automated log collection and analysis tools. Logs are written to the local system and then sent to the audit collection system for storage and review and analysis.

PRIVACY TESTING

Does the system require a System Test Plan?

No

Please explain why:

Testing is conducted for troubleshooting purposes and on an ad-hoc basis. While a specific test plan does not exist, the IBM team is following a defined testing process: associated with Security Technical Implementation Guides (STIGs) content published to the government checklist repository to determine compliance with system baseline requirements. These guides, when implemented, enhance security for software, hardware, physical and logical architectures to further reduce vulnerabilities. System Administrators (SA) are provided with validation instructions (Check Test) for the tested controls as well as corrective action steps ensuring remediation (Fix Text). Finding Details as well as Comments are added as evidence to the saved checklists by the SA To test and harden the system.

SBU DATA USE

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

Yes

Was permission granted per the requirements of Form 14664, SBU Data Use Questionnaire or Form 14665, SBU Data Use Request?

Yes

Provide the date the permission was granted.

7/30/2021

Was testing performed in conformance with IRM 10.8.8 Information Technology (IT) Security, Sensitive But Unclassified (SBU) Data Policy?

Yes

NUMBER AND CATEGORY OF PII RECORDS

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

IRS Employees: Not Applicable

Contractors: Not Applicable

Members of the Public: 100,000 to 1,000,000

Other: No

CIVIL LIBERTIES

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

Yes

Explain the First Amendment information being collected and how it is used.

Charitable contribution information is needed to figure and collect the right amount of tax on form 1040 US Individual Income Tax Return and attachment Schedule A, itemized deductions.

Please list all 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).

There is a statute that expressly authorizes its collection (identified in Q6).

Will the First Amendment information be used as the basis to make any adverse determination about an individual's rights, benefits, and/or privileges under Federal programs?

No

Is the system information used to conduct 'data-mining' as defined in the Implementing Recommendations of the 9/11 Commission Act of 2007, Public Law 110-53, Section 804?

No

Will this system have the capability to identify, locate, and monitor individuals or groups of people?

No

Does computer matching occur?

No

ACCOUNTING OF DISCLOSURES

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?

Yes

Does the system have a process in place to account for such disclosures in compliance with IRC §6103(p) (3) (A) or Subsection (c) of the Privacy Act? Contact Disclosure to determine if an accounting is required.

Not Applicable