

Date of Approval: 07/03/2024
Questionnaire Number: 1358

Basic Information/Executive Summary

What is the name of your project (system, database, pilot, product, survey, social media site, etc.)?

Robotic Process Automation Digitalization Enablement Platform (DEP) Form
Data Extraction

Acronym:

RPA DEP Form Data Extraction

Business Unit

Enterprise Digitalization

Preparer

For Official Use Only

Subject Matter Expert

For Official Use Only

Program Manager

For Official Use Only

Designated Executive Representative

For Official Use Only

Executive Sponsor

For Official Use Only

Executive Summary: Provide a clear and concise description of your project and how it will allow the IRS to achieve its mission.

This PCLIA applies to the Specialty Collection Offers in Compromise (SCOIC) Data Extraction automation. Offers in Compromise (OIC's) are currently being processed and stored in paper format, regardless of intake format. The solution seeks to automate the process of ingesting, processing, and storing of OIC documents. This automated solution includes the Optical Character Recognition/Intelligent Character Recognition-based extraction of critical fields from OIC documents and the transmission of the extracted data and images into an As-Received Data Storage. This solution creates a scalable solution for extracting data from digitized paper forms, and over 5 years, the automation will be used to save an estimated 170,000+ FTE hours.

Personally Identifiable Information (PII)

Will this project use, collect, receive, display, store, maintain, or disseminate any type of Sensitive but Unclassified (SBU), Personally Identifiable Information (PII), or Federal Tax Information (FTI)?

Yes

Please explain in detail how this project uses sensitive data from inception to destruction (data lifecycle).

The solution processes summary sheets from the AOIC system which contain PII data by extracting critical fields from the forms and transmit the data into a database

Please select all types of Sensitive but Unclassified data (SBU)/Personally Identifiable Information (PII)/Federal Tax Information (FTI) that this project uses.

Address

Employer Identification Number

Individual Taxpayer Identification Number (ITIN)

Name

Telephone Numbers

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

PII for federal tax administration - generally IRC Sections 6001 6011 or 6012

Product Information (Questions)

1.1 Is this PCLIA a result of the Inflation Reduction Act (IRA)?

Yes

1.2 What is the IRA Initiative Number?

Objective 1 Initiative 1.2

1.3 What type of project is this (system, project, application, database, pilot/proof of concept, power platform/visualization tool)?

Robotic Process Automation project

1.35 Is there a data dictionary for this system?

No

1.36 Explain in detail how PII and SBU data flow into, through and out of this system.

The UiPath automation will retrieve data from individual Automated Offer in Compromise (AOIC) summary sheets for Optical Character Recognition (OCR) extractions. The outputs from the OCR extraction will go through schema

validation. Once the process is complete, the validated extracted data will go to As-Received database.

1.4 Is this a new system?

Yes

1.5 Is there a Privacy and Civil Liberties Impact Assessment (PCLIA) for this system?

No

1.8 If the system is on the As-Built-Architecture, what is the ABA ID of the system? If this PCLIA covers multiple applications shown on the ABA, please indicate the ABA ID for each application covered separated by a comma.

Not Available

1.9 What OneSDLC State is the system in (Allocation, Readiness, Execution)?

Execution

2.1 If this system discloses any PII to any third party outside the IRS, 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. Enter "Yes" or "No". If Exempt, type "Exempt".

Exempt. The system does not disclose any PII to any third party outside the IRS.

2.2 Please provide the full name of and acronym of the governance board or Executive Steering Committee (ESC) this system reports to.

Internal Management Domain Governance Board (IMD GB)

3.1 Does your project/system involve any use of artificial intelligence (AI), including virtual assistant, chat bot, and robotic process automation, as defined in Executive Order 13960?

Yes

3.2 What is the methodology used and what database is training your AI?

Robotic Process Automation and it is being trained in UiPath.

3.3 Does this system use cloud computing?

Yes

3.31 Please identify the Cloud Service Provider (CSP), FedRAMP Package ID, and date of FedRAMP authorization.

Cloud Service Provider is Amazon Web Services. FedRAMP Package ID is F1603047866, FedRAMP authorization date is 6/21/2026

3.32 Does the CSP allow auditing?

Yes

3.32 Who has access to the CSP audit data (IRS or 3rd party)?

IRS

3.33 Please indicate the background check level required for the CSP (None, Low, Moderate or High).

Moderate

3.4 Is there a breach/incident plan on file?

Yes

3.5 Does the data physically reside in systems located in the United States and its territories and is all access and support of this system performed from within the United States and its territories?

Yes, for Amazon Web Services

3.6 Does this system interact with the public through a web interface?

No

3.61 If the system requires the user to authenticate, was a Digital Identity Risk Assessment (DIRA) conducted?

No

3.7 Describe the business process allowing an individual to access or correct their information.

The automation that resides on the Platform will ensure due process being provided to taxpayers, however the Platform by does not provide these capabilities. Notice, consent, and due process are provided in the tax forms instructions filed by the taxpayer, and pursuant to 5 USC.

4.1 Who owns and operates the system (IRS Owned and Operated, IRS Owned and Contractor Operated, Contractor Owned and Operated)?

IRS owned and operated

4.2 If a contractor owns or operates the system, does the contractor use subcontractors?

No

4.5 Identify the roles and their access level to the PII data. For contractors, indicate whether their background investigation is complete or not.

IRS Users - Read and Write access

IRS Managers - Read and Write access

IRS System Admin - Administrator

Background investigation for users and contractors involved is complete.

Contractors will not be using the system in Production as they will only be supporting the development of the tool.

4.51 How many records in the system are attributable to IRS Employees? Enter “Under 50,000”, “50,000 to 100,000”, “More than 100,000” or “Not Applicable”.

Not Applicable

4.52 How many records in the system are attributable to contractors? Enter “Under 5,000”, “5,000 to 10,000”, “More than 10,000” or “Not Applicable”.

Not Applicable

4.53 How many records in the system are attributable to members of the public? Enter “Under 5,000”, “5,000 to 10,000”, “More than 10,000” or “Not applicable”.

More than 10,000

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

The automation that resides on the Platform will access enterprise systems, however the Platform by does not have access to the enterprise systems. Assessment will be made on individual automation scripts to determine the access to SBU/PII data, and the corresponding systems owner will then determine and grant access to end user. Enterprise tools such as Business Entitlement Access Request System (BEARS) and established processes will be leveraged for granting access to the system.

5.1 Please describe any privacy risks, civil liberties and/or security risks identified for the system that need to be resolved and what is the mitigation plan?

No privacy risks, civil liberties or security risks have been identified for the system that require a mitigation plan.

5.11 Is there a Risk Assessment Form and Tool (RAFT) associated with this system on file with your organization or the IRS Risk Office.

No

5.2 Does this system use or plan to use SBU data in a non-production environment?

No

Interfaces

Interface Type

IRS Systems, file, or database

Agency Name

AO Offers (AOIC Summary Sheet)

Incoming/Outgoing

Incoming (Receiving)

Transfer Method

Electronic File Transfer Utility (EFTU)

Interface Type

IRS Systems, file, or database

Agency Name

S3 bucket

Incoming/Outgoing

Both

Transfer Method

Electronic File Transfer Utility (EFTU)

Interface Type

Forms

Agency Name

AOIC Summary Sheet

Incoming/Outgoing

Both

Transfer Method

Electronic File Transfer Utility (EFTU)

Interface Type

IRS Systems, file, or database

Agency Name

As Received Database

Incoming/Outgoing

Incoming (Receiving)

Transfer Method

Integrated Enterprise Portal (IEP)

Systems of Records Notices (SORNs)

SORN Number & Name

IRS 34.037 - Audit Trail and Security Records

Describe the IRS use and relevance of this SORN.

To identify and track any unauthorized accesses to sensitive and classified information and potential breaches. The use of records is or IRS to process AO Offers and not for any other unauthorized accesses as followed by the SORN

Records Retention

What is the Record Schedule System?

General Record Schedule (GRS)

What is the retention series title?

GRS 3.1 item 010

What is the GRS/RCS Item Number?

GRS 3.1 item 010

What type of Records is this for?

Both (Paper and Electronic)

Please provide a brief description of the chosen GRS or RCS item.

GRS 3.1 item 010-Infrastructure project records. Destroy 5 years after project is terminated, but longer retention is authorized if required for business use

What is the disposition schedule?

The retention period will be determined by the source systems policies and procedures

Data Locations

What type of site is this?

System

What is the name of the System?

As Received database

What is the sensitivity of the System?

Personally Identifiable Information (PII) including Linkable Data

Please provide a brief description of the System.

Storage and staging of data

What are the incoming connections to this System?

UiPath automation performs OCR extraction on AO Offers that contain the summary sheet retrieved from S3 bucket. The output of the extraction will go through schema validation, after that the validated extracted data will go to As-Received database

What are the outgoing connections from this System?

As Received database is the final location where the SCOIC extracted data will go to. Data can be retrieved via the Digital Content Retrieval (DCR) system and a summary report will be generated indicating downstream routing transactions.