

Date of Approval: 06/12/2024
Questionnaire Number: 1389

Basic Information/Executive Summary

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

CDW Knowledge Graph Environment

Acronym:

CKGE

Business Unit

Research Applied Analytics and Statistics

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.

CKGE is a data product that provides users with Compliance Data Warehouse (CDW) selected data information represented in a network/graph form, which provides more efficient means to identify relationships between tax entities.

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).

PII data from CDW, RRP, and LBI DataMart sources are extracted, transformed, and loaded into data stores (e.g., graph data and Elastic Search) that are used to provide the information available to entitled users. End users can query, view, and export networked data for their authorized use. Stored data is updated monthly, and processing year data deleted for any records older than 16 years.

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

Address

Centralized Authorization File (CAF)

Criminal Investigation Information

Criminal Record

Document Locator Number (DLN)

Employer Identification Number

Employment Information

Federal Tax Information (FTI)

Financial Account Number

Individual Taxpayer Identification Number (ITIN)

Internet Protocol Address (IP Address)

Passport Number

Preparer Taxpayer Identification Number (PTIN)

Professional License Number

Social Security Number (including masked or last four digits)

Tax ID Number

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

SSN for tax returns and return information - IRC section 6109

Product Information (Questions)

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

No

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

Data product for visualization.

1.35 Is there a data dictionary for this system?

Yes

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

For the persistent data, the data is queried by authorized staff from CDW and transformed and loaded into CKGE graph database/Elastic Search data stores. The data can then be visualized or queried by authorized users based on their approved uses. Data can be exported to laptops as Excel workbooks by users for use in their case files. Other queried data can be used by data scientists to develop model variables for their use projects.

1.4 Is this a new system?

No

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

Yes

1.6 What is the PCLIA number?

CKGE, PCLIA # 6150, approved 9/23/2021

1.7 What are the changes and why?

Regular renewal

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.

211385

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

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

CDAO DMD Change Control Board; CDAO Portfolio Review Board

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?

No

3.3 Does this system use cloud computing?

No

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

No

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

IRS Owned with support from contractors.

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.

All employees and contractors who administer and use the data product have MBI clearances, Executive approvals, and BEARS entitlements.

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".

Tax information for all IRS employees might be present.

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".

Tax information for all IRS contractors might be present.

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".

Tax information for all business and individual taxpayers might be present.

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

An approval form signed by an authorized executive must be attached with the BEARS entitlement request, which is then approved by 1) managers, 2) IT RAAS Reviewer Group, and 3) final account setup and approval by DMD.

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?

Privacy Risk: Taxpayer information of a limited amount that involved some taxpayer attributes related to filing and identifiers, along with relationships between taxpayers would be possible disclosed if a bad actor extracted the information. The information within CKGE is not complete tax filing information. Mitigation plan involves 1) access and event logs to Cybersecurity, and 2) Providing user's manager with service to review their employee's access activities.

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.

Yes

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

Compliance Data Warehouse

Incoming/Outgoing

Incoming (Receiving)

Transfer Method

Secured channel via HTTPS

Interface Type

IRS Systems, file, or database

Agency Name

LB&I DataMart

Incoming/Outgoing

Incoming (Receiving)

Transfer Method

Secured channel via HTTPS

Interface Type

IRS Systems, file, or database

Agency Name

RRP

Incoming/Outgoing

Incoming (Receiving)

Transfer Method

Secure File Transfer Protocol (SFTP)

Systems of Records Notices (SORNs)

SORN Number & Name

IRS 42.021 - Compliance Programs and Projects Files

Describe the IRS use and relevance of this SORN.

End users of CKGE are using the information for compliance programs.

Records Retention

What is the Record Schedule System?

Record Control Schedule (RCS)

What is the retention series title?

Schedule 27

What is the GRS/RCS Item Number?

27

What type of Records is this for?

Electronic

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

These disposition instructions will be included in IRS Document 12990 under Records Control Schedule 27 for Compliance Research, item 53 when next updated/published. Records are retained for this long to capture any recycling of tax evasion strategies for flow through-based tax schemes. Periodically, untoward actors will recycle schemes, and we have found that keeping past evidence assists in new investigations/audits. Often, the same individuals are involved.

What is the disposition schedule?

CKGE data is approved for destruction 16 years after processing year or when no longer needed for operational purposes, whichever is later (Job No. N1-58-09-86, approved 4/22/10).

Data Locations

What type of site is this?

System

What is the name of the System?

RAS-1

What is the sensitivity of the System?

Personally Identifiable Information (PII) including Linkable Data

What is the URL of the item, if applicable?

<https://cdw.web.irs.gov/Tools/environments/ckge/CkgeMain.aspx>

Please provide a brief description of the System.

RAS-1 is the DMD environment located within the IRS network boundary and further located behind a specific IRS firewall.

What are the incoming connections to this System?

IRS Data feeds and IRS workstation connections