

PRIVACY IMPACT ASSESSMENT

(Rev. 2/2020)

(All Previous Editions Obsolete)

Please submit your responses to your Liaison Privacy Official.

All entries must be Times New Roman, 12pt, and start on the next line.

http://intranet.epa.gov/privacy/pdf/lpo_roster.pdf. If you need further assistance, contact your LPO.

System Name:	
Renaming system to ARMS Uploader (previously called Content Ingestion Services)	
Preparer:	Office:
Security Point of Contact: Greg Adams/ Information System Security Officer: Nannette Willis	OMS/ORASE/ERMCD/CRTBORASEERMCD
Date:	Phone:
3/18/2025	919-541-4297/ 443-889-7402
Reason for Submittal: New PIA___ Revised PIA__X__ Annual Review___ Rescindment ___	
This system is in the following life cycle stage(s):	
Definition <input type="checkbox"/> Development/Acquisition <input type="checkbox"/> Implementation <input type="checkbox"/>	
Operation & Maintenance <input checked="" type="checkbox"/> Rescindment/Decommissioned <input type="checkbox"/>	
Note: New and Existing Systems require a PIA annually, when there is a significant modification to the system or where privacy risk has increased to the system. For examples of significant modifications, see <u>OMB Circular A-130, Appendix 1, Section (c) (1) (a-f)</u>.	
The PIA must describe the risk associated with that action. For assistance in applying privacy risk see <u>OMB Circular No. A-123, Section VII (A) (pgs. 44-45)</u>.	

Provide a general description/overview and purpose of the system:

The United States Environmental Protection Agency’s Enterprise Records Management Division (ERMCD) within the Office of Mission Support, Office of Records, Administrative Systems, and eDiscovery (ORASE)

provides leadership and direction in managing records that support EPA's mission. ERMCD, in collaboration with other Agency records management experts, is engaged in development, operation and maintenance of the ARMS Uploader, previously referred to as Content Ingestion Services, so as to provide better Records Management capabilities to EPA.

The main purpose of ARMS Uploader is to help users to submit records to the repository - in accordance with relevant EPA guidelines, by suggesting which record schedules may be appropriate for their documents. In addition to facilitate proper transfer of Records/Contents to the Repository, ARMS Uploader will support the centralized records digitization centers enhancements in order to ensure that sending records to ARMS from third party systems is as easy as possible. These enhancements include the development of a series of web services that can facilitate proper upload of records by ARMS.

The following web services will be made available:

- Text Extraction – This service will accept a variety of document formats and extract text from those documents.
- PDF OCR – This service will OCR scanned PDFs producing a searchable PDF.
- Text Summarization – This service will summarize text within a file.
- Records Schedule Categorization – Using a machine learning model, this service will return a suggested records schedule based on text of a file.

Above mentioned web services enable ARMS Uploader to provide its users record schedule suggestions as well as pass-through services to assist with Records Management. ARMS Uploader does not explicitly ask users for any privacy data and it does not store records but sends those records that it processes to ARMS.

Note:-

ARMS Uploader does not collect information. ARMS Uploader is a series of web services that analyse the content of records to provide record schedule suggestions. CIS provides pass through services, data are not stored by the four ARMS Uploader services described above. Data will be stored within ARMS central repository.

Section 1.0 Authorities and Other Requirements

1.1 What specific legal authorities and/or Executive Order(s) permit and define the collection of information by the system in question?

- Title III of the E-Government Act of 2002, Federal Information Security Management Act (FISMA) 44 U.S.C 3541, et seq
- The Privacy Act of 1974, PL 93-579, as amended 5 U.S.C 552a
- The Freedom of Information Act, PL 93-5025 U.S.C 552
- The Federal Managers' Financial Integrity Act (FMFIA), PL 97-25531 U.S.C 66a
- OMB Circular A-130, *Management of Federal Information Resources*
- OMB Circular A-123 Revised, *Management's Responsibility for Internal Control*, December 2004
- OMB Circular A-127, *Financial Management Systems*, July 23, 1993
- OMB Circular A-130, Appendix III, *Security of Federal Automated Information Resources*
- FIPS 200, *Minimum Security Requirements for Federal Information and Information Systems*
- FIPS 199, *Standards for Security Categorization of Federal Information and Information Systems*
- NIST SP 800-18, Revision 1, *Guide for Developing Security Plans for Federal Information Systems*
- NIST SP 800-53, Revision 4, *Recommended Security Controls for Federal Information Systems and Organizations*

1.2 Has a system security plan been completed for the information system(s) supporting the system? Does the system have or will the system be issued an Authorization-to-Operate? When does the ATO expire?

ARMS Uploader, previously referred to as CIS, is covered under a separate ATU currently. There are plans in place to merge the ARMS Uploader ATU under the existing ARMS ATO SSP. Consolidating ARMS Uploader and ARMS Nuxeo under a single ATO makes sense since these applications are part of the same records management system. Additionally, there is now an updated PIA that consolidates ARMS Uploader, ARMS Nuxeo, and the Paper Asset Tracking Tool (PATT).

1.3 If the information is covered by the Paperwork Reduction Act (PRA), provide the OMB Control number and the agency number for the collection. If there are multiple forms, include a list in an appendix.

No ICR required.

1.4 Will the data be maintained or stored in a Cloud? If so, is the Cloud Service Provider (CSP) FedRamp approved? What type of service (PaaS, IaaS, SaaS, etc.) will the CSP provide?

Yes, ARMS Uploader collects files submitted by end-users which is stored within a MongoDB SaaS database. The Mongo DB Atlas for Government is FedRAMP approved. Please see the following link for additional details on FedRAMP Moderate Authorization: <https://www.mongodb.com/products/platform/trust/fedramp> .

Section 2.0 Characterization of the Information

The following questions are intended to define the scope of the information requested and/or collected, as well as reasons for its collection.

2.1 Identify the information the system collects, uses, disseminates, or maintains (e.g., data elements, including name, address, DOB, SSN).

ARMS Uploader collects the following metadata associated with files end-users submit through the web interface: Title, Description, Sensitivity, Custodian, EPA Contacts, Record Schedule, record Close Date, Subjects, EPA Identifiers, Spatial Extent, Temporal Extent, and Tags. None of the metadata collected contains sensitive information. The metadata is not stored within ARMS Uploader, but passed to the ARMS Nuxeo system and ultimately stored within MongoDB.

2.2 What are the sources of the information and how is the information collected for the system?

The source of information are EPA emails or EPA documents that are determined to be records by EPA staff. These records pass through the ARMS Uploader and a record schedule prediction is provided to the end user.

2.3 Does the system use information from commercial sources or publicly available data? If so, explain why and how this information is used.

No.

2.4 Discuss how accuracy of the data is ensured.

Each individual employee is responsible for the content of their records. The role of CIS is to help users select the appropriate record schedule when they submit records to the ARMS records repository.

2.5 Privacy Impact Analysis: Related to Characterization of the Information

Discuss the privacy risks identified for the specific data elements and for each risk explain how it was mitigated. Specific risks may be inherent in the sources or methods of collection, or the quality or quantity of information included.

Privacy Risk:

CIS does not store any data – just provides pass-through service, the risk here is misuse of the data by internal users, another (potential) risk may be intentional/unintentional alteration of data.

Mitigation:

Privacy controls from NIST 800-53 are deployed on ARMS Uploader to ensure that any PII

contained in records passing through the application are not compromised. These include access controls (user group access controls, requirement of specific authorization for using sensitive data), restriction on any public release of the data, maintenance of audit, security of the data going through API gateway and monitoring of data usage (and periodically review/update approved procedures for data-handling if needed).

Section 3.0 Access and Data Retention by the System

The following questions are intended to outline the access controls for the system and how long the system retains the information after the initial collection.

3.1 Do the systems have access control levels within the system to prevent authorized users from accessing information they don't have a need to know? If so, what control levels have been put in place? If no controls are in place why have they been omitted?

All ARMS Uploader 3rd party application connection web services are called through an API management tool that will control the access to the web services. Access control levels are in place within ARMS to prevent unauthorized access to records that are analyzed by the 3rd party application connection web services. 3rd party application connections such as the ARMS Uploader end-user web interface must be registered through the API management tool and granted access to the 3rd party application connection web services.

3.2 In what policy/procedure are the access controls identified in 3.1, documented?

ARMS Uploader 3rd party application connections use an API management tool to control access to the web services. Access to the web services only allows a 3rd party application developer to post records to these services for analysis. ARMS which stores the actual records uses the following policy/procedure:

<https://usepa.sharepoint.com/sites/oei/ermd/ECMS/SitePages/Policy.aspx>

Access control will be documented in the 3rd Party Application Connections security documentation – in the spreadsheet Application Specific Security Controls

3.3 Are there other components with assigned roles and responsibilities within the system?

No, ARMS Uploader does not contain additional components with assigned roles and responsibilities within the system.

3.4 Who (internal and external parties) will have access to the data/information in the system? If contractors, are the appropriate Federal Acquisition Regulation (FAR) clauses included in the contract?

Only internal EPA 3rd party applications and staff will have access to ARMS Uploader but will not be able to access any data/information from the system because ARMS Uploader

does not store any data. Also, ARMS Uploader does not explicitly request or prompt end users to enter privacy data. The appropriate FAR clauses are included in the contract.

3.5 Explain how long and for what reasons the information is retained. Does the system have an EPA Records Control Schedule? If so, provide the schedule number.

ARMS Uploader does not store any data. All data are stored within ARMS central repository. Records that pass-through ARMS Uploader are retained in ARMS until scheduled disposition of records from the system based on pre-defined records retention schedules.

Schedule: 0742 - Agency Records Management System (ARMS)

3.6 Privacy Impact Analysis: Related to Retention

Discuss the risks associated with the length of time data is retained. How were those risks mitigated? The schedule should align the stated purpose and mission of the system.

Privacy Risk:

Records that pass-through ARMS Uploader must be retained in ARMS based on their pre-defined records retention schedules. Removal of records prior to the pre-defined records retention schedule is a potential privacy risk.

Mitigation:

Removal events are audited on a weekly basis by reports generated from the ARMS system. All deletions from the system are audited. The addition of records to the system is audited through weekly transaction reports and daily records transmission reports. The reports are maintained by the NCC ARMS system administrator and the EPA ARMS Program Management Office.

Section 4.0 Information Sharing

The following questions are intended to describe the scope of the system information sharing external to the Agency. External sharing encompasses sharing with other federal, state and local government, and third-party private sector entities.

4.1 Is information shared outside of EPA as part of the normal agency operations? If so, identify the organization(s), how the information is accessed and how it is to be used, and any agreements that apply.

No.

4.2 Describe how the external sharing is compatible with the original purposes of the collection.

No external sharing is done by using the ARMS Uploader.

4.3 How does the system review and approve information sharing agreements, MOUs, new uses of the information, new access to the system by organizations within EPA and outside?

The ARMS Uploader Application does not have an MOU/ISA with any other system.

4.4 Does the agreement place limitations on re-dissemination?

The ARMS Uploader Application does not have an agreement with any external system.

4.5 Privacy Impact Analysis: Related to Information Sharing

*Discuss the privacy risks associated with the sharing of information outside of the agency.
How were those risks mitigated?*

Privacy Risk:

None. No information sharing is done by the ARMS Uploader.

Mitigation:

None. The only information that is produced by ARMS Uploader is a record schedule suggestion. There is no risk of sharing of information within ARMS Uploader or to other 3rd party applications.

Section 5.0 Auditing and Accountability

The following questions are intended to describe technical and policy-based safeguards and security measures.

5.1 How does the system ensure that the information is used as stated in Section 6.1?

ARMS Uploader does not use the information but assists EPA employees. (All transactions are tracked within ARMS Uploader through auditable events using the API management tool. The API management tool does not alter the original content or time ordering of audit records. The API management tool handles audit reduction and report generation capability and audit logs are reviewed to ensure that information is used in accordance with approved practices.)

5.2 Describe what privacy training is provided to users either generally or specifically relevant to the system/collection.

Only EPA employees and approved contractors may request access to ARMS Uploader web services by registering their application, which means they have completed and passed Information Security and Privacy Awareness Training.

5.3 Privacy Impact Analysis: Related to Auditing and Accountability

Privacy Risk:

Audit logs could be inadvertently lost.

Mitigation:

Database audit logs are kept for 90 days, some of that time on the instance and some of the time archived on a separate instance. Anything older than 90 days is deleted daily. In general they are on the DB server for about 24 hours then moved to the backup appliance. Audit records generated today would stay on the database for the rest of the day, then be moved to the separate database and kept for 90 days. The last backup of the separate database would be kept for another 90 days. That adds up to approximately 180 days' retention, 170 days to allow for the usual "expected" unexpected.

Section 6.0 Uses of the Information

The following questions require a clear description of the system's use of information.

6.1 Describe how and why the system uses the information.

The ARMS Uploader application does not use the information within itself but assists EPA employees by suggesting which record schedule may be appropriate for their records.

6.2 How is the system designed to retrieve information by the user? Will it be retrieved by personal identifier? Yes X No . If yes, what identifier(s) will be used. (A personal identifier is a name, social security number or other identifying symbol assigned to an individual, i.e. any identifier unique to an individual. Or any identifier that can be linked or is linkable to an individual.)

ARMS Uploader retrieves agency user names which is required for implementing access controls within ARMS Nuxeo when end-users attempt to access submitted records.

6.3 What type of evaluation has been conducted on the probable or potential effect of the privacy of individuals whose information is maintained in the system of records?

[The goal here is to look at the data collected, how you plan to use it, and to ensure that you have limited the access to the people who have a need to know

in the performance of their official duties. What controls have you erected around the data, so that privacy is not invaded? ex. administrative control, physical control, technical control.]

Privacy threshold analysis was conducted. This system is also subject to privacy impact assessments such as the current one.

6.4 Privacy Impact Analysis: Related to the Uses of Information

Describe any types of controls that may be in place to ensure that information is handled in accordance with the uses described above.

Privacy Risk:

ARMS Uploader users can misuse the information in violation of the approved procedures and instructions.

Mitigation:

In addition to core controls per FISMA directive, *following technical controls are in place for tracking misuse*): Utilization of API security tool (APIMAN) which manages security of the data going through API gateway, User group access controls - requirement of specific authorization for use of data deemed sensitive by EPA, restriction (per EPA instructions) on any public release of the data, maintenance of audit logs and monitoring of data usage (including periodic review/update of approved procedures for data-handling if needed).

***If no SORN is required, STOP HERE.**

The NPP will determine if a SORN is required. If so, additional sections will be required.

Section 7.0 Notice

The following questions seek information about the system's notice to the individual about the information collected, the right to consent to uses of information, and the right to decline to provide information.

7.1 How does the system provide individuals notice prior to the collection of information? If notice is not provided, explain why not.

7.2 What opportunities are available for individuals to consent to uses, decline to provide information, or opt out of the collection or sharing of their information?

7.3 Privacy Impact Analysis: Related to Notice

Discuss how the notice provided corresponds to the purpose of the project and the stated uses. Discuss how the notice given for the initial collection is consistent with the stated use(s) of the information. Describe how the project has mitigated the risks associated with potentially insufficient notice and opportunity to decline or consent.

Privacy Risk:

Mitigation:

Section 8.0 Redress

The following questions seek information about processes in place for individuals to seek redress which may include access to records about themselves, ensuring the accuracy of the information collected about them, and/or filing complaints.

8.1 What are the procedures that allow individuals to access their information?

8.2 What procedures are in place to allow the subject individual to correct inaccurate or erroneous information?

8.3 Privacy Impact Analysis: Related to Redress

Discuss what, if any, redress program the project provides beyond the access and correction afforded under the Privacy Act and FOIA.

Privacy Risk:

Mitigation: