Affects Members Of the Public? ☒

Department of Energy

Privacy Impact Assessment (PIA)

MODULE I – PRIVACY NEEDS ASSESSMENT

Date	May 6, 2021	
Departmental Element & Site	Office of Inspector General, Tech Ops/Cyber Investigations & Forensic Analysis (CIFA)	
Name of Information System or IT Project	CDS – Federal Cloud Discovery Services (CDS)	
Exhibit Project UID	019-000002764	
New PIA ☒ **Update** ☐	New PIA.	
	Name, Title	**Contact Information Phone, Email**
System Owner	John Pizzurro Special Agent in Charge Tech OPS/CIFA	John.Pizzurro@hq.doe.gov 202-586-0111
Local Privacy Act Officer	Alexander Borman	1000 Independence Ave. SW Washington DC 20585 202-586-1653 Alexander.borman@hq.doe.gov
Cyber Security Expert reviewing this	Jennifer Purcell	202-586-7136

MODULE I – PRIVACY NEEDS ASSESSMENT		
document (e.g. ISSM, CSSM, ISSO, etc.)		Jennifer.purcell@hq.doe.gov
Person Completing this Document	Sarah Nelson	1000 Independence Ave. SW Washington DC 20585 202-586-1668 Sarah.nelson@hq.doe.gov
Purpose of Information System or IT Project	Complete Discovery Source's Cloud Electronic Discovery Services (CDS) Software-as-a-Service (SaaS) platform is a secure solution for meeting electronic discovery needs. The eDiscovery application Relativity is offered as a single Commercial Off-The-Shelf (CotS) service by CDS. CDS has built a segregated physical and logical environment for its FedRAMP CDS clients. CDS runs in fully redundant Windows server environments and includes content analysis of underlying proprietary databases with a web-based graphical user interface (GUI), which consists of clustered VMware servers. All client data is stored on servers running in an active/active cluster configuration. CDS utilizes a series of redundant Cisco firewalls, routers, and switches to manage network traffic. Production equipment is hosted at the Equinix, Inc. data center located in North Bergen, New Jersey. The failover datacenter is located in Washington D.C. The purpose of the information system is to collect digital information in a user-friendly environment in order to provide direct electronic discovery support for all components of the OIG including, but not limited to audits, inspections, data analytics, and investigations. For example, the information system may be used to organize the extensive electronic evidence gathered during a criminal investigation. Additionally, the information system may be used to facilitate both external and internal litigation. For example, the information system may be used in conjunction with the OIG's Office of Counsel to facilitate the discovery process in matters before the Merit Systems Protection Board and Equal Employment Opportunity Commission. The system will also be evaluated for its usefulness in responding to Freedom of Information Act requests directed to the OIG. With today's data-driven environment, this information system will play a critical role in supporting the OIG's core mission of strengthening the integrity, economy, and efficiency of the Department's programs and operations. In-scope information is collected by DOE personnel, and then "processed" into the system. Processing, involves parsing out meta-data, performing any Optical Character Recognition (OCR) to ensure data is searchable, and filtering data by dates, file types, Custodians, etc. Performing these actions enables the system to index data for	

MODULE I – PRIVACY NEEDS ASSESSMENT	
	discovery and enables ease of retrieval through the search function. Data, after processing, is organized as it was when provided to CDS, i.e., all the emails for a custodian, which were processed from their PST file, will appear in the same directory structure (inbox, Outbox, Sent, and any other folders that the user created). This structure is not static, however, and can be modified in advance or on the fly to meet the needs of OIG personnel. Custodian files can be organized in many ways including source structure, or by any other means; 1) data type, 2) key words, 3) concepts, 4) search terms, etc. Data retrieval is performed via OIG personnel searches as well as the subsequent export of those search results for use during trial or investigation. Users can query the system for any type of information, performing simple one-word searches, or complex searches that combine fielded data with key terms, people, and dates. Within this retrieval processes users looking for PII/PHI/PCI can query this information based on patterns of anonymized data, or tags. When querying the system, users looking for a specific piece of PII/PHI/PCI can use the exact set of numbers, e.g. 123-45-6789. Documents brought back will contain that exact set of digits. Users may also use pattern recognition to retrieve all document, in this case, with a social security number, simply by searching for "XXX-XX-XXXX". The system recognizes this input and will retrieve all documents with this pattern on information embedded. All queries can be further limited through the application of filters, like thoughts run during the processing phase (dates, people, terms, organizations, concepts, etc.). Relativity supports reporting on any information contained within the system. Users can structure reports for documents based on any field of information. Additionally, the system maintains compliance reports governing users and administrative functions. All of these can be run ad-hoc, or on pre-scheduled basis.
Type of Information Collected or Maintained by the System:	☒ SSN ☒ Medical & Health Information ☒ Financial Information ☒ Clearance Information ☐ Biometric Information ☒ Mother's Maiden Name

MODULE I – PRIVACY NEEDS ASSESSMENT

	☒ DoB, Place of Birth ☒ Employment Information ☒ Criminal History ☒ Name, Phone, Address ☒ Other – Contracts, Acquisition Information, Regulations, Laws, etc.
Has there been any attempt to verify PII does not exist on the system? **DOE Order 206.1, *Department of Energy Privacy Program,* defines PII as** *any information collected or maintained by the Department about an individual, including but not limited to, education, financial transactions, medical history and criminal or employment history, and information that can be used to distinguish or trace an individual's identity, such as his/her name, Social Security number, date and place of birth, mother's maiden name, biometric data, and including any other personal information that is linked or linkable to a specific individual.*	PII exists on the system.
If "Yes," what method was used to verify the system did not contain PII? (e.g. system scan)	N/A
Threshold Questions	
1. Does system contain (collect and/or maintain), or plan to contain any information about individuals?	Yes
2. Is the information in identifiable form?	Yes
3. Is the information about individual Members of the Public?	Yes
4. Is the information about DOE or contractor employees?	☒ Federal Employees ☒ Contractor Employees

MODULE I – PRIVACY NEEDS ASSESSMENT

END OF PRIVACY NEEDS ASSESSMENT

MODULE II – PII SYSTEMS & PROJECTS

AUTHORITY, IMPACT & NOTICE

1. AUTHORITY

What specific authorities authorize this system or project, and the associated collection, use, and/or retention of personal information?

- The Privacy Act, 5 U.S.C. § 552a, authorizes each agency to maintain information about an individual as is relevant and necessary to accomplish a purpose of the agency as required by statute or by Executive Order of the President.

- Public Law (PL) 95-452, the Inspector General Act of 1978, as amended [Title 5 United States Code (U.S.C.), App. 3], sets forth authority and functions of the Inspector General and provides authorization to collect and maintain personal information to uncover fraud, waste, abuse, misuse, corruption, criminal acts, or mismanagement relating to DOE programs, operations, facilities, contracts, and information technology systems.

- In accordance with DOE O 221.2A, Cooperation with the Office of Inspector General, during the courses of audits, inspections, or investigations all DOE employees and contractors must:

  1. Cooperate fully and promptly with requests from the OIG for information and data relating to DOE programs and operations;

  2. Comply with requests for interviews and briefings and must provide affidavits or sworn statements, if so requested by an employee of the OIG so designated to take affidavits or sworn statements; and

  3. Not impede or hinder other employees' cooperation with the OIG.

- Additionally, in accordance with DOE O 221.2B, Reporting Fraud, Waste and Abuse to the Office of Inspector General, employees and contractors must:

  1. Report allegations and/or actual instances of fraud, waste, abuse, misuse, corruption, criminal acts, or mismanagement relating to DOE programs to the OIG; and

  2. Supervisors must not impede, discourage, or prohibit

	employees from reporting such matters directly to the OIG.
2. CONSENT **What opportunities do individuals have to decline to provide information (e.g. where providing information is voluntary) or to consent only to particular uses of the information (other than required or authorized uses)?**	To effectuate the OIG's mission under the Inspector General Act of 1978, as amended, opportunities are not provided to individuals wishing to withhold or place limitations on investigatory information in the system. This information system is exempt from certain requirements under 10 CFR § 1008.12(a), which provides for an exemption for *Investigative Files of the Inspector General (DOE-54)*. Moreover, 10 CFR § 1008.12(b)(2) provides a broad exemption for investigatory material compiled for law enforcement purposes. These exemptions enable the OIG to avoid notifying an individual at the individual's request of the existence of records in an investigative file pertaining to such individual or granting access to an investigative file in certain circumstances.
3. CONTRACTS **Are contractors involved with the design, development and maintenance of the system? If yes, was the Privacy Order CRD or Privacy Act clauses included in their contracts?**	Yes. The contractor being utilized for the design and development of the system is Complete Discovery Source's Cloud Electronic Discovery Services (CDS). Relativity is offered as a single Commercial Off-The-Shelf (COTS) service. Two Privacy Clauses are in the contract: (1) (58) 52.239-1, Privacy or Security Safeguards (AUG 1996) (5 U.S.C. 552a) and (2) (xx)(A) 52.224-3, Privacy Training (JAN 2017) (5 U.S.C. 552a). CDS will be providing cloud hosting and administrative support but will have no direct access to the information in the software. There will be no OIG contractor personnel administering the system. Specifically, CDS will be handling the data sets provided by the OIG for the purpose of making it available (loading or processing) to the OIG within Relativity. CDS has full administrative access within the software of Relativity but will not be analyzing or reviewing substantive content within the system. Post loading/processing, CDS will assist the OIG with functionality related requests such as searches, setting up email threading, or coding layouts as well as matter-related requests such as bates stamping (organizational method for labeling and identifying legal documents) images for a production set when requested by OIG personnel.
4. IMPACT ANALYSIS: **How does this project or information system impact privacy?**	If PII within Relativity is compromised, it would have a serious adverse effect on individuals in light of the highly sensitive information the system will contain relating to investigations of, and litigation involving, individuals. Relativity may contain sensitive personal information (e.g., SSN, personal financial information, personal health information) as well as negative personal information (i.e., criminal and noncriminal investigatory information). Given the highly sensitive

	nature of documents acquired through discovery, a compromise of the information in this system could create significant harm and embarrassment for individuals including, but not related to, potentially significant reputational harm, professional harm, financial harm, embarrassment relating to the disclosure of personal health information, identity harm, and damage to trust between individuals and the Federal Government. The data in the system is relevant and necessary for the OIG to perform its statutory responsibilities. The OIG has mitigated the risk of a data breach by encrypting data. There are no connecting systems with direct access to the documents in the system, which will lower the risk of an indirect data breach via a different information system. Internally, the system will be strictly controlled, and data will only be maintained in the system the duration necessary to accomplish the necessary task. For example, with respect to litigation before the Equal Employment Opportunity Commission, the information will be removed from the system after the discovery process has been completed and the Office of Counsel has confirmed that it no longer has a need to access the information on the specific system. PII will not be maintained in the system past the expiration of a specified purpose.
5. SORNs **How will the data be retrieved? Can PII be retrieved by an identifier (e.g. name, unique number, or symbol)?** **If yes, explain, and list the identifiers that will be used to retrieve information on the individual.**	PII can be retrieved using Relativity. Relativity utilizes a software component called dtSearch. dtSearch is a basic program that indexes data and allows users to search through data using Boolean search identifiers such as words, numbers, and dates. Depending on the type of case the OIG is investigating, and the specific facts associated with that case, we intend to use specific words, names, dates, company names, contract numbers, and any other basic identifiers associated with the facts of the case to retrieve data. However, some documents may not support optical character recognition (OCR), and thus not support index searches. Consequently, although PII may be present in the documentation supporting an audit, inspection, investigation, and/or litigation, it may not be retrievable.
6. SORNs **Has a Privacy Act System of Records Notice (SORN) been published in the *Federal Register*?** **If "Yes," provide name of SORN and location in the *Federal Register*.**	• Investigative Files of the Inspector General (DOE-54) • Allegation-Based Inspection Files of the Office of Inspector General (DOE-83).

7. SORNs **If the information system is being modified, will the SORN(s) require amendment or revision?**	N/A
DATA SOURCES	
8. What are the sources of information about individuals in the information system or project?	Relativity may collect PII from a variety of potential sources including individuals and through the course of active OIG investigations. Depending on the size of the data, information is either uploaded by the individual directly into Relativity or sent to the vendor CDS if the file(s) are too large for CDS to upload directly to the server. The OIG acquires documentation in paper form as well as digital data and electronically stored information (ESI) by several means, including IG subpoenas, search warrants, consent searches, official data request memos, and complaint intake forms. The paper and digital data collected during official OIG investigations may contain PII including names, dates of birth, addresses, and social security numbers, etc.
9. Will the information system derive new or meta data about an individual from the information collected?	Relativity produces reports which may contain meta data including analysis or determinations relating to custodian data corresponding to individuals. This includes any long text fields where stakeholders have inputted analysis or determinations about custodian data/documents. These reports are not automated and must be run by either the user (with permission) or CDS Project Managers.
10. Are the data elements described in detail and documented?	Yes, the data elements are described within the Security documents of the A&A package.
DATA USE	
11. How will the PII be used?	PII collected is processed in support of an audit, inspection, investigation, and/or litigation. However, Relativity is only used temporarily in the life cycle of the data to process it. It is not used to house the data indefinitely or for any time period exceeding the need to support the conclusion reached in an OIG audit, inspection, investigation, and/or litigation.

12. If the system derives meta data, how will the new or meta data be used? **Will the new or meta data be part of an individual's record?**	Meta data in the form of analysis or determinations may be used to support a conclusion in an audit, inspection, investigation, or litigation.
13. With what other agencies or entities will an individual's information be shared?	Information may be viewed by another Federal agency's OIG during the triennial peer review process or during investigation work, such as a review by the Government Accountability Office (GAO) or Federal Bureau of Investigation (FBI).
REPORTS	
14. What kinds of reports are produced about individuals or contain an individual's data?	There are no automated reports produced about individuals. Only in the event an individual were subject of a review would a summary report be derived from that person's personal data. A summary report can include metadata and the user is able to export those fields directly from the platform in the form of a report. This includes any long text fields where stakeholders have inputted analysis or determinations about custodian data/ documents. These reports are not automated and must be run by either the user (with permission) or CDS Project Managers; however, all work product is saved in the event that reports need to be run. These work products include narrative/summary reports, audit reports, and other reports (i.e., litigation) that will be made available for OIG internal use.
15. What will be the use of these reports?	Only in the event an individual was the subject of a review, an auditor or inspector's narrative report (also known as a "summary report") would be used if such action was warranted.
16. Who will have access to these reports?	OIG users will have access to the data depending on the individual's need to access the system. However, the ability to access information on the system will be restricted on whether the OIG personnel are administrators, users, or reviewers. Each category of personnel will have different privileges. Only the System Admin and Case Admin will have access to narrative/summary reports as well as other reports. As noted, no automated reports will be produced. When creating reports derived from information in the system, the OIG carefully follows the requirements of both the Privacy Act and the Freedom of Information Act to ensure that PII is protected to the extent permitted by law.

Role	Service Provider or Internal	Sensitivity Level	Authorized Privileges and Functions Performed
System Admin	Service Provider	Moderate	User Creation, User Privilege Assignment, Case Creation
Case Admin	Internal	Moderate	Case Specific User Creation, Case Specific User Privilege Assignment
Case User	Internal	Low	Case Preparation
Basic Reviewer	Internal External	Low Low	Case Review, Case Research Case Review, Case Research

MONITORING

17. Will this information system provide the capability to identify, locate, and monitor individuals?

Unlikely, but some documents may include location information. Documents are compartmentalized upon processing within a case or matter. In addition, once data is ingested into a particular matter it can be further divided into a granular folder structure based on specific search terms, custodians, date ranges etc. Once in the database, documents can be searched and tagged accordingly. There are fields that contain information like custodian name, date last modified, to, from, cc, etc. and these can be customized to reflect case specific details. Should a user manually tag a document according to location data (e.g., an individual's address), this may provide the capability to locate that individual, but this would be a rare use of the system. There are no automated processes that would facilitate such a function.

18. What kinds of information are collected as a function of the monitoring of individuals?

N/A

19. Are controls implemented to prevent unauthorized monitoring of individuals?

Yes, information that resides in the system is in an encrypted form and there is no formal or automated method of monitoring an individual's information that may have been collected during field work. All data entering the CDS network is subject to anti-virus and anti-malware scans using TrendMicro software. Suspicious files are immediately quarantined, and a notification email is sent to the IT department for additional actions. In addition, sensitive data is compartmentalized and protected by role-based access controls.

Role	Service Provider or Internal	Sensitivity Level	Authorized Privileges and Functions Performed
System Admin	Service Provider	Moderate	User Creation, User Privilege Assignment,

			Case Creation
Case Admin	Internal	Moderate	Case Specific User Creation, Case Specific User Privilege Assignment
Case User	Internal	Low	Case Preparation
Basic Reviewer	Internal External	Low Low	Case Review, Case Research Case Review, Case Research

DATA MANAGEMENT & MAINTENANCE

20. How will records about individuals be kept current and verified for accuracy, relevance, and completeness? Include PII data collected from sources other than DOE records.

The records will be maintained in accordance with generally accepted Government audit standards and the standards of the Council of the Inspectors General on Integrity and Efficiency (CIGIE). Additionally, investigators will review all information gathered. Any conclusions will be corroborated with sufficient and reliable sources to ensure that the data is complete, accurate, relevant, and balanced. Additionally, prior to publication (the formulation of an investigation report), all investigative, audit, and inspection reports as well as their supporting documents are reviewed by multiple levels of management and verified by an independent reviewer during a factual accuracy verification process.

21. If the information system is operated in more than one site, how will consistent use of the information be ensured at all sites?

Client data and specific case data (per client request) are segregated into dedicated network shares. There are no directly connecting information systems with access to the documents in the system. Access and functional permissions are specific to the share and the user's role. High level functions are limited to administrative users and access to the database is limited to the database administrators. Authentication of internal users is run against the Active Directory using PIV credentials per DOE/EITS permission, which has already been granted. Users who are at the Savannah River and Lexington sites will have to use a soft token provided by CDS. The soft token app will be downloaded to the user's DOE-issued cell phone.

Data enters the system in two ways, both of which are protected by appropriate data security measures:

1. Delivery of encrypted physical devices to CDS facilities;
2. Via Secure File Transfer Protocol (SFTP). All data uploaded via the CDS SFTP is protected with SSL encryption while in transit.

All physical devices are affixed with a unique identifier upon receipt. This identifier is used to track physical chain-of-custody of the device. A chain-of-custody log is maintained with entries being added with each change of custody. Data from physical devices are forensically

	transferred to the CDS network for staging. The term "forensically" refers to the discipline of digital forensics. Data that is forensically sound is data in digital format that has been acquired, transferred, processed, and analyzed in a way that preserves the data in its original form and does not alter the data in any way. Data is handled in a forensically sound manner to include maintaining a strict chain of custody to ensure the integrity of the data for admissibility into civil and criminal court proceedings. All data entering the CDS network is subject to anti-virus and anti-malware scans using TrendMicro software. Suspicious files are immediately quarantined, and a notification email is sent to the IT department for additional actions. An MD5 hash & Secure Hash Algorithm (SHA) value is generated for each file as it is ingested into the CDS system and is used throughout the process to verify its authenticity. Forensically identical copies of files are used for all the processes in the CDS system while unadulterated original files are maintained.
RECORDS MANAGEMENT	
22. Identify the record(s).	1. Semiannual Reports to Congress 2. Audit Records 3. Inspection Records 4. Investigative Records
23. Identify the specific disposition authority (is) that correspond to the record(s) noted in no. 22.	☒ Scheduled ▪ Semiannual Reports to Congress. (Ni-434-00-1, Item 1) ▪ Audit Records. (N1-434-00-1, Item 2) ▪ Inspection Records. (Ni-434-00-1, Item 3) ▪ Investigative Records. (N1-434-00-1, Item 4)
24. Records Contact	Alexander Borman (Records Liaison Officer) 202-586-1653 Alexander.borman@hq.doe.gov
ACCESS, SAFEGUARDS & SECURITY	
25. What controls are in place to protect the data from unauthorized access, modification, or use?	Management, technical, and administrative controls are in place to protect the data from unauthorized access, modification, or use. Two factor authentication is required for access. Remote access is available only through DOEnet through a VPN. After a connection has been established to DOEnet, users authenticate to the CDS server using a client application installed on DOE-owned laptops with full disk encryption.

	All physical devices are affixed with a unique identifier upon receipt. This identifier is used to track physical chain-of-custody of the device. A chain-of-custody log is maintained with entries being added with each change of custody. Data from physical devices are forensically transferred to the CDS network for staging. All data entering the CDS network is subject to anti-virus and anti-malware scans using TrendMicro software. Suspicious files are immediately quarantined, and a notification email is sent to the IT department for additional actions. An MD5 hash & Secure Hash Algorithm (SHA) value is generated for each file as it is ingested into the CDS system and is used throughout the process to verify its authenticity. Forensically identical copies of files are used for all the processes in the CDS system while unadulterated original files are maintained.
26. Who will have access to PII data?	Access is limited to authorized OIG investigators, investigative analysts, auditors, attorneys, and inspectors. Additionally, access may be made available to Management whose official duties require access to the records. Access is controlled on a per project basis which allows only assigned team members with a need-to-know access to a particular audit or inspection. Other Federal agencies may be granted read-only access during the triennial peer review process or if a relevant audit, inspection, or investigation is being conducted by an entity such as GAO or FBI.
27. How is access to PII data determined?	Users' access is controlled by the system administrator on a per-project basis which allows assigned staff limited ability to access information relating to matters to which they are assigned.
28. Do other information systems share data or have access to the data in the system? If yes, explain.	No.
29. For connecting information systems, is there an Interconnection Security Agreement (ISA) or other agreement between System Owners to ensure the privacy of individuals is protected?	N/A

30. Who is responsible for ensuring the authorized use of personal information?	The CDS System Owner and the System Administrators.
END OF MODULE II	

SIGNATURE PAGE		
	Signature	**Date**
System Owner	**John Pizzurro** John Pizzurro Digitally signed by John Pizzurro Date: 2021.07.14 12:36:09 -04'00' **(Signature)**	
Local Privacy Act Officer	**Alexander Borman** Alexander J. Borman Digitally signed by Alexander J. Borman Date: 2021.07.14 12:43:16 -04'00' **(Signature)**	
Ken Hunt **Chief Privacy Officer**	**(Print Name)** William K. Hunt Digitally signed by William K. Hunt Date: 2021.07.14 11:58:17 -04'00' **(Signature)**	