Privacy Impact Assessment (PIA)
for the

ED Cyber Data Lake
November 30, 2021

For PIA Certification Updates Only: This PIA was reviewed on **Enter date** by **Name of reviewer** certifying the information contained here is valid and up to date.

Contact Point

Contact Person/Title: Roman Kulbashny Branch Chief, Security Engineering and Architecture Branch
Contact Email: Roman.Kulbashny@ed.gov

System Owner

Name/Title: Roman Kulbashny Branch Chief, Security Engineering and Architecture Branch
Principal Office: Office of the Chief Information Officer (OCIO)

Please submit completed Privacy Impact Assessments to the Privacy Office at privacysafeguards@ed.gov

*Please complete this **Privacy Impact Assessment (PIA)** on how personally identifiable information (PII) is collected, stored, protected, shared, and managed electronically by your system. You may wish to consult with your ISSO in completing this document.*
If a question does not apply to your system, please answer with N/A.

1. Introduction

1.1. Describe the system including the name, acronym, and a brief description of the program or purpose for the system.

ED Cyber Data Lake (EDCDL) will be the repository for all actionable cyber operation, risk management, information security continuous monitoring, continuous diagnostic and mitigation, and other data relevant to the U.S. Department of Education (Department) Cybersecurity Program. EDCDL is a data lake, which is a repository designed to store large amounts of source system data in native form. These data can be structured, semi-structured, or unstructured data. EDCDL is a dependent system that relies on the Department Splunk Cloud (EDSplunk) as a cloud service provider to store the data. EDCDL will act as the Department Security Information and Event Management security platform, allowing authorized Department personnel to view the security events that are collected by Department information technology (IT) systems as part of normal operations and monitoring. EDCDL will provide strictly maintained and monitored role-based access to collected security events data for the Department's dedicated cybersecurity personnel. These personnel will monitor IT systems for compliance with cybersecurity policy and standards and detecting cybersecurity incidents, including advanced threat detection and near real-time monitoring for indicators of compromise of IT systems that have been integrated with EDCDL.

EDCDL will receive the previously collected actionable security events data from the Department's information systems, along with the associated time stamp (i.e., the time of the event that was logged) and other systems' attributes for each data record in order to analyze the event and assess risks posed to the Department. Security event data are not retrieved by name or other personal identifier but rather by time stamp of the event, criticality of the event, associated vulnerability, hardware asset information, and other information systems attributes. Information will not be retrieved from this system using PII, and no PII will disseminated outside of the system. Only Department cybersecurity personnel who are fully cleared and explicitly assigned with responsibilities to monitor for cybersecurity incidents and security posture of IT systems will have access to security events

collected by EDCDL. The information collected by EDCDL will be used by cleared cybersecurity personnel to maintain integrity, availability, and confidentiality of Department IT systems and information.

The data received by EDCDL from Department IT systems are categorized as: security management, information security, IT infrastructure maintenance, record retention, system maintenance, contingency planning, enterprise architecture, inspections, and auditing. Examples of information types include: vulnerabilities identified on systems, cybersecurity threat information, firewall logs, network access lists, systems logs, information about running process, information about IT assets (e.g., operating systems, host name, location, organizations ownership), information about software installed on system (e.g., name of vendor, version of the software), records about planned and performed systems management activities (e.g., time when system has been started, shutdown, security or system patches have been installed, missing patches). In some cases, PII such as Social Security numbers (SSNs) and credit card numbers can be received incidentally from source systems in the process of monitoring. If a source system detects and collects PII during systems and security monitoring and does not sanitize this PII in its systems and security logs, EDCDL may receive PII that is collected by a source system, although this is rare. As a result, this PII received by EDCDL will be stored with associated records in accordance with record management schedule of associated records, with a retention period that is normally one or three years, but will not be used to identify the affected individual(s). For example, EDCDL can receive PII from the Department's Data Loss Prevention tool if PII is sent from the Department email system outside of the Department without being encrypted. Another example: a Department system detects a phishing or ransomware attack, creates a security event record containing associated PII (e.g., name of individual or email address) and sends this security event to EDCDL.

1.2. Describe the purpose for which the personally identifiable information (PII)[1] is collected, used, maintained or shared.

EDCDL does not intentionally collect PII. However, PII may be received incidentally by EDCDL from other Department systems that have appropriate authorization to collect and use PII. EDCDL will receive the previously collected actionable security events data from the Department's information systems with the associated time stamp for each data record. Time stamps will be the essential data element to search and access security event data. Only Department cybersecurity personnel and

[1] The term "personally identifiable information" refers to information which can be used to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc. OMB Circular A-130, page 33

contractors who are fully cleared and explicitly assigned with responsibilities to monitor for cybersecurity incidents and security posture of IT systems will have access to security event data collected by EDCDL. The information collected by EDCDL will be used by cybersecurity personnel to maintain integrity, availability, and confidentiality of Department IT systems and information. In some cases, certain pieces of PII (such as SSNs and credit card numbers) may be received incidentally from source systems in the process of monitoring for compliance with Department "Standard PR.DS: PII Data Loss Prevention (DLP)". For example, EDCDL may incidentally receive the PII detected by the Department's DLP system. If a source system detects and collects PII during systems and security monitoring and does not sanitize this PII in its systems and security logs, EDCDL may receive PII that is collected by a source system. As a result, this PII will be received by EDCDL but will not be used to identify the affected individual(s), is not used in the security analysis process, and is never shared by EDCDL with other systems. All PII is encrypted at rest and in transfer by default in EDCDL system boundaries.

1.3. Is this a new system, or one that is currently in operation?

New System

1.4. Is this PIA new, or is it updating a previous version?

New PIA

1.5. Is the system operated by the agency or by a contractor?

Agency

1.5.1. If the system is operated by a contractor, does the contract or other acquisition-related documents include privacy requirements?

☑ N/A

Click here to select.

2. Legal Authorities and Other Requirements

If you are unsure of your legal authority, please contact your program attorney.

2.1. What specific legal authorities and/or agreements permit and regulate the collection and use of data by the system? Please include name and citation of the authority.

In accordance with the Federal Information Security Modernization Act of 2014 (FISMA), the Department has established an Enterprise-wide Information Security Program (ISP) to safeguard the confidentiality, integrity, and availability of its information and systems. PII incidentally received by EDCDL is initially collected in compliance with the Department's overarching Cybersecurity Policy OCIO-3-112, to analyze, identify, alert and prevent the unintentional or deliberate exfiltration of unprotected sensitive data from the Department's network. The Cybersecurity Policy is based on legislative, statutory, and executive directive requirements that include Federal laws and regulations, Presidential Directives and Executive Orders, Federal IT Acquisition Reform Act (FITARA), FISMA, the National Institute of Standards and Technology (NIST) Special Publications (SP) 800 series, the NIST Federal Information Processing Standards (FIPS), and Office of Management and Budget (OMB) memoranda. The Department has an obligation under Federal law to define and operate an effective cybersecurity program.

At a minimum, this requirement is driven by:

Privacy Act of 1974, 5 U.S.C. § 552a(e)(10). Agencies are required to "establish appropriate administrative, technical, and physical safeguards to ensure the security and confidentiality of records and to protect against any anticipated threats or hazards to their security or integrity which could result in substantial harm, embarrassment, inconvenience or unfairness to any individual on whom information is maintained."

FISMA, Public Law No: 113-283, directs agency heads to ensure that: (1) information security management processes are integrated with budgetary planning; (2) senior agency officials, including chief information officers, carry out their information security responsibilities; and (3) all personnel are held accountable for complying with the agency-wide information security program.

> § 3553. Authority and functions of the Director and the Secretary, Part (a), Subpart (2) requires agencies: "to identify and provide information security protections commensurate with the risk and magnitude of the harm resulting from the unauthorized access, use, disclosure, disruption, modification, or destruction of —
>
> (A) Information collected or maintained by or on behalf of an agency; or
>
> (B) Information systems used or operated by an agency or by a contractor of an agency or other organization on behalf of an agency".

§ 3554. Federal agency responsibilities, Part (a), Subpart (7) requires: "(b) AGENCY PROGRAM. — Each agency shall develop, document, and implement an agency-wide information security program to provide information security for the information and information systems that support the operations and assets of the agency, including those provided or managed by another agency, contractor, or other source."

SORN

2.2. Is the information in this system retrieved by an individual's name or personal identifier such as a Social Security Number or other identification?

No

2.2.1. If the above answer is **YES,** this system will need to be covered by Privacy Act System of Records Notice(s) (SORN(s)).[2] Please provide the SORN name, number, Federal Register citation and link, or indicate that a SORN is in progress.

☑ N/A

Click here to enter text.

2.2.2. If the above answer is **NO**, explain why a SORN was not necessary. For example, the information is not retrieved by an identifier, the information is not maintained in a system of records, or the information is not maintained by the Department, etc.

☐ N/A

The information is not retrieved by a PII identifier, the information is not maintained in any system of record as it is collected incidentally in the collection of other security information, the PII data received by EDCDL is not used in the security analysis.

Records Management

If you do not know your records schedule, please consult with your records liaison or send an email to RMHelp@ed.gov

2.3. What is the records retention schedule approved by National Archives and Records Administration (NARA) for the records contained in this system? Please provide all relevant NARA schedule numbers and disposition instructions.

[2] A System of Records Notice (SORN) is a formal notice to the public that identifies the purpose for which PII is collected, from whom and what type of PII is collected, how the PII is shared externally (routine uses), and how to access and correct any PII maintained by ED. https://connected.ed.gov/om/Documents/SORN-Process.pdf

General Records Schedule (GRS) 3.2, Item 010 Systems and data security records.
Temporary. Destroy 1 year(s) after system is superseded by a new iteration or when no longer needed for agency/IT administrative purposes to ensure a continuity of security controls throughout the life of the system.
DAA-GRS2013-0006- 0001
General Records Schedule (GRS) 3.2, 020 Computer security incident handling, reporting and follow-up records
Temporary. Destroy 3 year(s) after all necessary follow-up actions have been completed, but longer retention is authorized if required for business use.
DAA-GRS2013-0006- 0002

2.4. Is the PII contained in this system disposed of appropriately, and in accordance with the timelines in the records disposition schedule?

Yes

3. Characterization and Use of Information

Collection

3.1. List the specific PII elements (e.g., name, email, address, phone number, date of birth, Social Security, etc.) that the system collects, uses, disseminates, or maintains.

The system does not intentionally collect PII. Information such as SSN, name, credit card number, email address, and username can be received from other Department systems that collect and process PII. If security event information includes PII, it can be sent to EDCDL. For example, EDCDL can receive information that a password for a username has been compromised. This security event can include username and/or name of employee.

The system will collect email addresses, usernames, and passwords from Federal employees to establish access credentials.

3.2. Does the system collect only the minimum amount required to achieve the purpose stated in Question 1.2?

Yes

EDCDL will receive the previously collected actionable security events data from the Department's information systems, along with the associated time stamp (i.e., the time

of the event that was logged) and other systems' attributes for each data record in order to analyze the event and assess risks posed to the Department. In some cases, PII such as SSNs and credit card numbers can be received incidentally from source systems in the process of monitoring. If a source system detects and collects PII during systems and security monitoring and does not sanitize this PII in its systems and security logs, EDCDL may receive PII that is collected by a source system, although this is rare. As a result, this PII received by EDCDL will be stored with associated records in accordance with record management schedule of associated records.

3.3. What are the sources of PII collected (e.g., individual, school, another agency, commercial sources, etc.)?

PII is received incidentally by EDCDL from other Department systems that have appropriateauthorization to collect and use PII.

3.4. How is the PII collected from the stated sources listed in Question 3.3 (e.g., paper form, web page, database, etc.)?

The information is electronically collected from other Department systems as necessary and stored in raw data form encrypted in EDCDL.

3.5. How is the PII validated or confirmed to ensure the integrity of the information collected?[3] Is there a frequency at which there are continuous checks to ensure the PII remains valid and accurate?

Since the PII is collected incidentally and it is not used for security analysis purposes, itis not validated. Please refer to the PIAs for other Department systems for information on how the PII is validated in those systems.

Use

3.6. Describe how the PII is used to achieve the purpose stated in Question 1.2 above.

PII is collected incidentally and not used for security analysis purposes.

3.7. Is the system using PII for testing/researching new applications or information systems prior to deployment or for training employees?

[3] Examples include restricted form filling, account verification, editing and validating information as it's collected, and communication with the individual whose information it is.

No

3.7.1. If the above answer is **YES,** what controls are in place to minimize the risk and protect the data?

☑ N/A

Click here to enter text.

Social Security Numbers

It is the Department's Policy that, in order to collect Social Security Numbers, the System Owner must state the collection is: 1) authorized by law, 2) necessary for an agency purpose, and 3) there is no reasonable alternative.

3.8. Does the system collect Social Security Numbers? Note that if the system maintains Social Security Numbers but does not explicitly collect them, answer 3.8.1 to address the purpose for maintaining them.

Yes

3.8.1. If the above answer is **YES**, explain the purpose for its collection, and how the SSN will be used.

☐ N/A

SSNs are collected incidentally in support of departmental DLP policy and are not used for identifying individuals. SSNs will be preserved only as part of raw data received by EDCDL from other ED systems and are encrypted at rest (See section 1.2 for more detail on information collection by EDCDL).

3.8.2. Specify any alternatives considered in the collection of SSNs and why the alternatives were not selected.

☑ N/A

Click here to enter text.

4. Notice

4.1. How does the system provide individuals with notice about the collection of PII prior to its collection (e.g., direct notice, such as a Privacy Act Statement (if applicable) or public notice, such as a SORN, PIA,)? If notice is not provided, explain why not.

Notice is provided to individuals by the other Department systems that collect the PII. Since the collection of PII in EDCDL is incidental and the PII is not used for security analysis, no additional notice is provided.

4.2. Provide the text of the notice or the link to the webpage where the notice is posted if notice is provided other than by SORN or PIA.

☑ N/A

Please refer to the PIAs for other Department systems for information about privacy notice.

4.3. What opportunities are available for individuals to consent to uses (including new uses of previously collected PII), decline to provide PII, or opt out of the project?

Please refer to the PIAs for other Department systems for information about opportunities for consent.

4.4. Is the notice referenced in Question 4.1 reviewed and revised when there are changes in the practice, policy, or activities that affect the PII and privacy to ensure that individuals are aware of and can consent to, where feasible, these changes?

N/A

5. Information Sharing and Disclosures

Internal

5.1. Will PII be shared internally with other ED principal offices? If the answer is **NO**, please skip to Question 5.4.

No

5.2. What PII will be shared and with whom?

☑ N/A

Click here to enter text.

5.3. What is the purpose for sharing the specified PII with the specified internal organizations?

☑ N/A

Click here to enter text.

External

5.4. Will the PII contained in the system be shared with external entities (e.g. another agency, school district, the public, etc.)? If the answer is **NO**, please skip to Question 6.1.

No

5.5. What PII will be shared and with whom? List programmatic disclosures only.[4]
Note: If you are sharing Social Security Numbers externally, please specify to whom and for what purpose.

☑ N/A

Click here to enter text.

5.6. What is the purpose for sharing the PII with the specified external entities?

☑ N/A

Click here to enter text.

5.7. Is the sharing with the external entities authorized?

☑ N/A

Click here to select.

5.8. Is the system able to provide and retain an account of any disclosures made and make it available upon request?

☑ N/A

Click here to select.

5.9. How is the PII shared with the external entity (e.g. email, computer match, encrypted line, etc.)?

☑ N/A

Click here to enter text.

5.10. Is the sharing pursuant to a Computer Matching Agreement (CMA), Memorandum of Understanding (MOU), or other type of approved sharing agreement with another agency?

☑ N/A

[4] If this information is covered by Privacy Act System of Records Notice (SORN) please list only relevant programmatic disclosures listed under the Routine Uses section.

Click here to select.

5.11. Does the project place limitation on re-disclosure?

☑ N/A

Click here to select.

6. Redress

6.1. What are the procedures that allow individuals to access their own information?

Individuals do not have access to their own information in EDCDL. However, individuals may have access to the information in other Department systems. Please refer to the PIAs for other Department systems for information about access.

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

PII is incidentally collected as part security event monitoring. Please refer to the PIAs for other Department systems for information about correction and amendment.

6.3. How does the project notify individuals about the procedures for correcting their information?

Please refer to the PIAs for other Department systems for information about correction and amendment.

7. Safeguards

If you are unsure which safeguards will apply, please consult with your ISSO.

7.1. Does the principal office work with their CSO/ISSO to build privacy & security into the system and build privacy extensions to the extent feasible?

Yes

7.2. Is an Authority to Operate (ATO) required?

Yes

7.3. Under NIST FIPS Pub. 199, what is the security categorization of the system: **Low, Moderate, or High?**

☐ N/A

Moderate

7.4. What administrative, technical, and physical safeguards are in place to protect the information?

EDCDL is hosted in a secure FedRAMP environment in the Amazon Web Services (AWS) and IBM Smart Cloud for Government (SCG). FedRAMP-authorized CSPs for Federal agencies and customers is strictly configured in accordance with specific regulatory and compliance requirements.

Access to the Enterprise EDCDL system is available only to users who have been authenticated to the Department of Education network using their Department-issued PIV card. Access to all privileged roles is controlled through processes that enforce formal requests and approvals for access on a need to know and least privilege basis. Enhancing this model, strict separation of duties is in place as well with regards to the distribution of roles. Access to data is protected through physical access controls to the hosting facilities, firewalls, network and host intrusion detection systems, event monitoring systems, nightly backups, and data encryption while at rest and in transit. Additionally, there are scheduled system audits, user recertification and vulnerability scans.

Only authorized users are connected to EDCDL via a certified Managed Trusted Internet Protocol Services (MTIPS) connection which monitored by a security operation center (SOC). EDCDL is rated Moderate by the FIPS 199 Security Categorization. EDCDL received an authorization to operate in July 2021.

EDCDL uses Access Control Lists (ACLs), firewalls, Intrusion Protection Systems (IPS), FIPS 140 validated encryption, multi-factor authentication, antimalware, and multiple cybersecurity capabilities to protect the information. PII is encrypted in transfer and at rest within the system boundaries of EDCDL. These security measures limit data access to Department and contract staff on a "need to know" basis and control individual users' ability to access information within the system.

Finally, all privileged users are provided a copy of the Rules of Behavior and are required to complete the annual Cybersecurity and Privacy Awareness training.

7.5. Is the information in the system appropriately secured in accordance with the IT security requirements and procedures as required by Federal law and policy?

Yes

7.6. Has a risk assessment been conducted where appropriate security controls to protect against that risk have been identified and implemented?

Yes

7.7. Please describe any monitoring, testing or evaluation conducted on a regular basis to ensure the security controls continue to work properly at safeguarding the PII.

EDCDL is reviewed annually and as needed when significant changes to the system occur. As part of the Department's continuous monitoring program, EDCDL is expected to review and renew the authorization to operate on a regular basis via the ongoing assessment and authorization process. This process includes audits of the implemented security and privacy controls by independent assessors. Findings from these audits produce Plans of Actions and Milestones (POAMs) for the system owner to remediate. Self-assessments are also conducted on a continuous basis, including annual incident response and contingency plan testing. On a more frequent basis, vulnerability and compliance scans are performed to check for vulnerabilities and deviations from the Department standards.

8. Auditing and Accountability

8.1. How does the system owner assess and ensure that the PII is used in accordance with stated practices in this PIA?

The system owner ensures the EDCDL systems administrators complete reviews of audit logs on a regular basis to ensure there is no misuse or malicious activity with the system or data.

The system owner periodically reviews audit reports provided by the EDCDL administrators regarding information processing and maintains the access control list of who can access PII. The system owner also works directly with the Department's privacy office on privacy compliance documentation to ensure all information in this PIA is up to date and accurate. Ultimately, the EDCDL system will undergo annual OMB Circular A-123, Appendix A (Management's Responsibility for Enterprise Risk Management and Internal Control) assessment, and NIST SP 800-53 system security control self-assessments.

Finally, the system owner documents the privacy controls every two years through the PIA review process. These privacy controls are then assessed by the security authorization team and the privacy office.

8.2. Does the system owner continuously monitor and audit the privacy controls to ensure effective implementation?

Yes

8.3. What are the privacy risks associated with this system and how are those risks mitigated?

Risks to privacy include unauthorized access, as well as mishandling and misuse of the information maintained in EDCDL.

To mitigate these privacy risks, EDCDL operates a comprehensive security program over the entire system and its supporting business processes. A key component of the security program is the continuous monitoring effort which ensures that the security and privacy controls remain effective over-time and that new threats are assessed, and appropriate countermeasures implemented.

The risk of unauthorized access to PII is mitigated through an array of safeguards, including: strict access controls, segregation of duties, physical access controls at the hosting facility, data encryption (both in flight and at rest), annual access certifications, and network and host-based intrusion detection systems. All user access is approved via established Department Privilege User Access process. EDCDL access is authorized only from internal network connections, only using multi-factor authentication and only form government-furnished equipment that have full disk encryption. The user list is reviewed on regular basis, and users are removed when they leave the organization or change positions. There are controls in place to monitor, review, and prevent unusual and unauthorized EDCDL access.

The risk of mishandling or misuse of PII is mitigated through a series of requirements for all EDCDL system administrators and users:

- Both prior to employment and as part of continuous monitoring, EDCDL system administrators are subject to background checks;
- Prior to gaining system access with elevated privileges, system administrators and users are required to sign a Rules of Behavior which governs their actions; and