Privacy Impact Assessment (PIA): CDC - CDH - QTR3 - 2023 - CDC6795139

Date Approved: 1/24/2025

General Information

PIA Name:	CDC - CDH - QTR3 - 2023 - CDC6795139	**PIA ID:**	1679901
Name of Component:	CDC - NNDSS Mercury	**Name of ATO Boundary:**	NNDSS Mercury
Office:		**OPDIV:**	CDC

1:	Identify the Enterprise Performance Lifecycle Phase of the system.	Operations and Maintenance
2:	Is this a FISMA-Reportable system?	Yes
3:	Does the system have or is it covered by a Security Authorization to Operate (ATO)?	Yes
4:	ATO Date or Planned ATO Date.	5/31/2024
5:	Is the system or electronic information collection, agency or contractor operated?	Agency

PTA

PTA - 2:	Indicate the following reason(s) for this PTA. Choose from the following options.	PIA Validation (PIA Refresh)
PTA - 3:	Is the data contained in the system owned by the agency or contractor?	Agency
PTA - 4:	Please give a brief overview and purpose of the system by describing what the functions of the system are and how the system carries out those functions.	CSELS Data Hub (CDH) is designed to receive public health data feeds from Electronic Health Records (EHR) from various partners to support the COVID-19 data needs and build a repository of datasets relevant to situational awareness and for ongoing public health responses, both present and future.
PTA - 5:	List and/or describe all the types of information that are collected (into), maintained, and/or shared in the system regardless of whether that information is PII and how long that information is stored.	The data in CDH is provided by private and public organizations including healthcare agencies, universities, labs, and federal agencies. It is

		important to note that the CDH platform does not collect the data directly from users but receives large datasets collected from the healthcare agencies. CDH data types collected from facilities, labs, pharmacies, National Emergency Medical Services (EMS) Information System (NEMSIS), and patients consist of organizational structure, services, beds, utilization, staffing, expenses, physician arrangements, system affiliation, geographic indicators, accreditations, approval codes, hospital readiness, adoption level of computerized systems, decision support, specialty, National Provider Information (NPI) code, specimen data, enrollment data, inpatient and outpatient medical claims, outpatient pharmacy claims, event code, procedures done by Emergency Medical Technicians (EMT), medications administered by EMTs, Reason for call (if reported to 9-1-1), non-identifiable unique patient ID. Patient's PII data consists of name, driver's license number, passport numbers, mailing address including county and census tract, email addresses, phone number, occupation, medical notes including symptoms, test results, dates, and codings, medical records number, military status, contact tracing, date of birth, sex, race, ethnicity, marital status, citizenship, tribal affiliation, healthcare facilities where the patient appointments occur, Logical Observation Identifiers Names and Codes (LOINC), and interpreted results as well as numeric result values and reference ranges, results include both positive and negative test results. Educational institutions and other organizations may use an identifier to identify their own users (e.g., Student ID). External non-CDC user credentials are collected and stored in CDH and consists of username, email address, phone number, full name, and password. Internal CDC users accessing the system are identified and authenticated via CDC's Active Directory (AD); AD is a separate system with its own PIA.
PTA - 5A:	Are user credentials used to access the system?	Yes, but the user credentials are maintained in a separate system (e.g., AD, AMS) and not collected or maintained by this system. The system providing credentials is
PTA - 5B:	Please identify the type of user credentials used to access the system.	
PTA - 6:	Describe why all types of information is collected (into), maintained, and/or shared with another system. This description should specify what information is collected about each category of individual.	The CSELS Data Hub Platform works in collaboration with participating commercial labs, vendors having public health datasets, EHR vendors, public

		organizations, federal agencies and universities that have agreed to share data with proper data use agreement (DUA) with CDC. The system does not directly collect the data from individuals. The data will be used by the CDC system Data Collation and Integration for Public Health Event Response (DCIPHER), Health and Human Services (HHS) Protect (Cloud Palantir) system to enrich the emergency department data and other datasets being provided to platform by multiple federal agencies to create a national picture. The data will also be used by CDC to do the needed research for their mission, by Emergency Operations Center (EOC) to derive useful information needed for timely decision making, by CDC studies to answer public health questions, and perform exploratory analysis and visualization. The data will be stored in the FedRAMP Amazon Web Services (AWS) environment. The data will be shared securely using AWS secure file transfer protocol service, access to Simple Scalable Storage (S3) service and Application Programming Interface (API) by enforcing strict role based access control (RBAC) and only providing read only access to data. The data will be maintained as per the duration agreed upon in the DUAs. It will be purged after the agreed upon duration or moved to archival (Amazon Glacier) when the latest data arrive, and older data is no longer needed. External non-CDC user credentials are collected and stored in CDH and consists of username, email address, phone number, full name, and password. Internal CDC users accessing the system are identified and authenticated via CDC's Active Directory (AD); AD is a separate system with its own PIA.
PTA - 7:	Does the system collect, maintain, use or share PII?	Yes
PTA - 7A:	Does this include Sensitive PII as defined by HHS?	Yes
PTA - 8:	Does the system include a website or online application?	No
PTA - 8A:	Are any of the URLs listed accessible by the general public (to include publicly accessible log in and internet websites/online applications)?	
PTA - 9:	Describe the purpose of the website, who has access to it, and how users access the web site (via public URL, log in, etc.). Please address each element in your response.	
PTA - 10:	Does the website have a posted privacy notice?	
PTA - 11:	Does the website contain links to non-federal government websites external to HHS?	
PTA - 11A:	Is a disclaimer notice provided to users that follow external links to websites not owned or operated by HHS?	
PTA - 12:	Does the website use web measurement and customization technology?	
PTA - 12A:	Select the type(s) of website measurement and customization technologies in use and if it is used to collect PII.	
PTA - 13:	Does the website have any information or pages directed at children under the age of thirteen?	

PTA - 13A:	Does the website collect PII from children under the age thirteen?	
PTA - 13B:	Is there a unique privacy policy for the website and does the unique privacy policy address the process for obtaining parental consent if any information is collected?	
PTA - 14:	Does the system have a mobile application?	No
PTA - 14A:	Is the mobile application HHS developed and managed or a third-party application?	
PTA - 15:	Describe the purpose of the mobile application, who has access to it, and how users access it. Please address each element in your response.	
PTA - 16:	Does the mobile application/ have a privacy notice?	
PTA - 17:	Does the mobile application contain links to non-federal government websites external to HHS?	
PTA - 17A:	Is a disclaimer notice provided to users that follow external links to resources not owned or operated by HHS?	
PTA - 18:	Does the mobile application use measurement and customization technology?	
PTA - 18A:	Describe the type(s) of measurement and customization technologies or techniques in use and what information is collected.	
PTA - 19:	Does the mobile application have any information or pages directed at children under the age of thirteen?	
PTA - 19A:	Does the mobile application collect PII from children under the age thirteen?	
PTA - 19B:	Is there a unique privacy policy for the mobile application and does the unique privacy policy address the process for obtaining parental consent if any information is collected?	
PTA - 20:	Is there a third-party website or application (TPWA) associated with the system?	No
PTA - 21:	Does this system use artificial intelligence (AI) tools or technologies?	No

PIA		
PIA - 1:	Indicate the type(s) of personally identifiable information (PII) that the system will collect, maintain, or share.	Name Email Address Phone numbers Medical records (PHI) Military Status Date of Birth Mailing Address Medical Records Number Employment Status Passport Number User Credentials Patient ID Number Driver License Number

		Other - Free text Field - Race, Sex, Ethnicity, Citizenship, Tribal Affiliation; Occupation; marital status.
PIA - 2:	Indicate the categories of individuals about whom PII is collected, maintained or shared.	Patients Members of the public
PIA - 3:	Indicate the approximate number of individuals whose PII is maintained in the system.	Above 2000
PIA - 4:	For what primary purpose is the PII used?	The PII data is strictly used for research purposes only, disease surveillance and reporting
PIA - 5:	Describe any secondary uses for which the PII will be used (e.g. testing, training or research).	Stored user credentials and passwords are used to authenticate external users.
PIA - 6:	Describe the function of the SSN, Truncated SSN, and/or Taxpayer ID.	N/A Neither SSN nor Taxpayer ID are collected
PIA - 6A:	Cite the legal authority to use the SSN, Truncated SSN, and/or Taxpayer ID.	N/A SSN is not collected
PIA - 7:	Identify legal authorities governing information use and disclosure specific to the system and program.	Public Health Service Act, sec. 301, Research and Investigation (42 U.S.C. 241); secs. 304, 306, and 308(d), which discuss authority to grant assurances of confidentiality for health research and related activities (42 U.S.C. 242b, 242k, and 242m(d)); sec. 361, Quarantine and Inspection, Control of Communicable Diseases (42 U.S.C. 264); and sec. 361F-3, Public Readiness and Emergency Preparedness Act (42 U.S.C. 247d-6d).
PIA - 8:	Are records in the system retrieved by one or more PII data elements?	Yes
PIA - 8A:	Please specify which PII data elements are used to retrieve records.	Name Email Address Phone Numbers Medical Records (PHI) Race Sex Ethnicity Citizenship Tribal Affiliation
PIA - 8B:	Provide the number, title, and URL of the Privacy Act System of Records Notice (SORN) that is being used to cover the system or indicate whether a new or revised SORN is in development.	09-90-2001 Records Used for Surveillance and Study of Epidemics, Preventable Diseases and Problems
PIA - 9:	Identify the sources of PII in the system.	Government Sources Within the OPDIV Other HHS OPDIV State/Local/Tribal

		Other Federal Entities Non-Government Sources Private Sector
PIA - 10:	Is there an Office of Management and Budget (OMB) information collection approval number?	No
PIA - 10A:	Provide the information collection approval number.	
PIA - 10B:	Identify the OMB information collection approval number expiration date.	
PIA - 10C:	Explain why an OMB information collection approval number is not required.	System does not collect data directly from individuals but obtains records that were collected from other agencies.
PIA - 11:	Is the PII shared with other organizations outside the system's Operating Division?	Yes
PIA - 11A:	Identify with whom the PII is shared or disclosed.	Other Federal Agency/Agencies State or Local Agency/Agencies Within HHS
PIA - 11B:	Please provide the purpose(s) for the disclosures described in PIA - 11A.	Other Federal Agency/Agencies - The data is used for disease surveillance and reporting. The system provides data for analysis to associate disease trends among groups like people within a certain age bracket, gender, geographic location, nationality, or race in order to derive useful information for decision making by public health officials, influence policies and communicate critical information. State or Local Agency/Agencies The data is used for disease surveillance and reporting. The system provides data for analysis to associate disease trends among groups like people within a certain age bracket, gender, geographic location, nationality, or race in order to derive useful information for decision making by public health officials, influence policies and communicate critical information. Within HHS - The data is used for disease surveillance and reporting. The system provides data for analysis to associate disease trends among groups like people within a certain age bracket, gender, geographic location, nationality, or race in order to derive useful information for decision making by public health officials, influence policies and communicate critical information. CSELS Data Hub platform requires data use agreements (DUAs) between all entities that connect to the platform that governs the use of all the data including PII.

PIA - 11C:	List any agreements in place that authorizes the information sharing or disclosure (e.g., Computer Matching Agreement (CMA), Memorandum of Understanding (MOU), or Information Sharing Agreement (ISA)).	CSELS Data Hub platform requires data use agreements (DUAs) between all entities that connect to the platform that governs the use of all the data including PII. The HIPAA Privacy Rule expressly permits the sharing of PHI for specified public health purposes, to public health authorities (e.g., state/local departments of health) that are legally authorized to receive/collect the information for the purpose of preventing or controlling disease, injury, or disability. See 45 C.F.R. 164.512(b). These sharing situations do not require agreements.
PIA - 11D:	Describe process and procedures for logging/tracking/accounting for the sharing and/or disclosing of PII. If no process or procedures are in place, please explain why not.	Accounting for disclosures when they happen is achieved via email logs. Logs are maintained by the CDH system administrators.
PIA - 12:	Is the submission of PII by individuals voluntary or mandatory?	Voluntary
PIA - 12A:	If PII submission is mandatory, provide the specific legal requirement that requires individuals to provide information or face potential civil or criminal penalties.	
PIA - 13:	Describe the method for notifying individuals that their information will be collected and how they can opt-out of the collection or use of their PII. If there is no option to object to the information collection, provide a reason.	There is no option because CDH platform is the recipient of data that has already been collected by data providers through the established agreements, procedures and arrangement they traditionally have followed for data collection. Individuals requesting to opt-out must do so according to the policies and procedures, systems, and options provided by data providers to individuals.
PIA - 14:	Describe the process to notify and obtain consent from the individuals whose PII is in the system when major changes occur to the system (e.g., disclosure and/or data uses have changed since the notice at the time of original collection). Alternatively, describe why they cannot be notified or have their consent obtained.	The CDH platform does not have a process to obtain consent from or notify individuals about data collection and use as this system is not the original collection of the data. The data in CDH has already been collected by data providers through their agreements with healthcare facilities, participating organizations and individuals. Notification and obtaining consent from individuals about data use is the responsibility of the data providers that collect it.
PIA - 15:	Describe the process in place to resolve an individual's concerns when they believe their PII has been inappropriately obtained, used, or disclosed, or that the PII is inaccurate. If no process exists, explain why not.	The CDH platform does not have a process in place to work with individuals regarding concerns about their PII stored in the system. CDH does not collect data from individuals. Each

		contributor to the CDH signs a Data Use Agreement (DUA) that specifies if a breach is discovered the data provider will be notified within the hour allowing the data provider that ability to communicate with the individuals. The system does not receive username/password in any data provided by the data providers. The system uses AWS Cognito service integrated with AWS Transfer service that uses username/password for a data provider to send data to CDC. The password is managed by Cognito service and is not available for any admin of the system to see.
PIA - 16:	Describe the process in place for periodic reviews of PII contained in the system to ensure the data's integrity, availability, accuracy and relevancy. Please address each element in your response. If no processes are in place, explain why not.	It is the primary responsibility of the data provider to maintain integrity, availability, accuracy and relevancy of the data in the system of origin used for providing data to CDC. In CDH original data received from data provider is maintained in its raw form to meet the requirement, while copy of data is created for data processing, visualization, reporting and analytic need. The data provider may periodically provide composite data that includes all the data previously provided and new data. The version process in the system maintains the old version while making the latest data available to user, thus maintaining integrity and accuracy of data while making multiple versions available and option to revert to any previous version.
PIA - 17:	Identify who will have access to the PII in the system.	Users Administrators Developers Contractors
PIA - 17A:	Select the type of contractor.	HHS/OpDiv Direct Contractors
PIA - 17B:	Do contracts include Federal Acquisition Regulation (FAR) and other appropriate clauses ensuring adherence to privacy provisions and practices?	Yes
PIA - 18:	Provide the reason why each of the groups identified in PIA - 17 needs access to PII.	Users - By using shared data from multiple jurisdictions (shared per fully executed data-use

		agreements), agencies can perform disease analysis. Administrators - Administrators are required to have access to the database to maintain the system. Developers: Developers build the data engineering pipeline to ingest, process, and make the data available for analytics. They touch the data to understand the data types and nature of data and any special handling it requires to build correct data and analytics pipeline. Contractors - Direct Contractors will be users and administrators of the system depending on individual's role to either perform analysis or to maintain the system.
PIA - 19:	Describe the administrative procedures in place to determine which system users (administrators, developers, contractors, etc.) may access PII.	Least Privileges is part of the structure of CDH and each user is placed in an appropriate group which limits the data they are able to view.
PIA - 20:	Describe the technical methods in place to allow those with access to PII to only access the minimum amount of information necessary to perform their job.	The CDH systems uses the National Institute of Standards and Technology (NIST) SP-800-53 revision 4 control AC-6 Least Privilege as the foundation to define and limit access to PII and the minimum amount of information for users to perform their job. The user's supervisor defines the minimum information system access that is required in order for the user to complete his/her job. The access list for the information system is reviewed on a quarterly basis during which time users' access permissions are reviewed/adjusted, and unneeded accounts are purged from the system. The limited access for reader role is enforced using AWS Identity and Access Management (IAM), AWS Lake Formation, Cloud Tamer, and CDC Active Directory.
PIA - 21:	Identify the general security and privacy awareness training provided to system users (system owners, managers, operators, contractors and/or program managers) using the system to make them aware of their responsibilities for protecting the information being collected and maintained.	All CDC personnel are required to take annual Privacy and Security Awareness Training.
PIA - 22:	Describe the training system users receive (above and beyond general security and privacy awareness training).	None
PIA - 23:	Describe the process and guidelines in place with regard to the retention and destruction of PII. Cite specific National Archives and Records Administration (NARA) records retention schedule(s) and include the retention period(s).	CDH data is kept by the CDC as a historical public health record, per CDC's "Scientific and Research Project Records Control Schedule", section 1a ("Authorized Disposition: PERMANENT"). Records Schedule N1-442-09-1. The data will be securely purged from the system if the agreed upon DUA requires purging the data after the agreed upon duration.

PIA - 24:	Describe how the PII will be secured in the system using administrative, technical, and physical controls. Please address each element in your response.	Physical controls: CDH data is protected by physical controls including physical barriers and locked doors protecting restricted areas, security guards, video cameras, climate control (cooling), redundant power systems, and fire prevention and control systems. Technical controls: CDH technical controls include Role-Based Access Control for all users (system administrators, developers, and users), encryption, multiple firewalls, and system redundancy. The system also undergoes continuous monitoring using automated monitoring systems. Administrative controls: CDH administrative controls include restricted access to the system; each individual user is vetted by CDH program management. Security Awareness Training is required and must be updated annually. Background checks are required for all users. OMB (Office of Management and Budget), HHS and CDC security and privacy policies and standards are followed by all users of the system.