US Department of Health and Human Services

Privacy Impact Assessment

Date Signed:

08/19/2020

OPDIV:

NIH

Name:

National Covid Cohort Collaborative (N3C)

PIA Unique Identifier:

P-5515149-785938

The subject of this PIA is which of the following?

Minor Application (stand-alone)

Identify the Enterprise Performance Lifecycle Phase of the system.

Development

Is this a FISMA-Reportable system?

No

Does the system include a Website or online application available to and for the use of the general public?

Yes

Identify the operator.

Contractor

Is this a new or existing system?

New

Does the system have Security Authorization (SA)?

Yes

Describe the purpose of the system.

The National CoVid Cohort Collaborative (N3C) system will serve as a cloud-based hosting environment for multiple data models that are transformed into a common analytic model for performing cohort analysis of CoVid-19 datasets and ongoing research involving public health actions, clinical care, and treatment.

N3C system is hosted on the Palantir Foundry Cloud Service (PFC) platform, a Federal Risk and Authorization Management Program (FedRAMP) authorized cloud service.

N3C is a collaborative partnership between the National Center for Advancing Translational Sciences (NCATS), the National Patient-Centered Clinical Research Network (PCORnet), Observational Health Data Sciences and Informatics (OHDSI), Accrual to Clinical Trials\ Informatics for Integrating Biology and the Bedside (ACT/i2b2), TriNetX, and several HHS agencies.

Describe the type of information the system will collect, maintain (store), or share.

The N3C will resource CoVid-related clinical data from the following:
Limited data sets (LDS) which may include dates (such as admission, discharge, service), date of birth (DOB), date of death (DOD); city, state, five digit or more zip code, and age

De-identified data of medical records that include visits & procedure occurrences, diagnosis, diseases, prescription & drugs, medical devices and supplies, medical conditions, date of death and health care provider locations.

Synthetic data sets that includes an artificial, statistically-comparable, computational derivative of the original data.

These data sets are authorized through Data Transfer Agreements (DTA) & Data Sharing & User Agreements (DUA) titled the NIH CoVid-19 Data Warehouse.

N3C users log in using one of the following:
NIH Login - NIH Identity, Credential, and Access Management Services: Identity Management Services (IMS), which maintains its own unique privacy impact assessment (PIA) on record, including all legal authorities documented.

HHS Personal Identity Verification (PIV), a US Federal government wide credential used to access Federally controlled facilities and information systems at the appropriate security level

Login.gov -a publicly available secure online access to participating government programs

Provide an overview of the system and describe the information it will collect, maintain (store), or share, either permanently or temporarily.

The N3C will resource CoVid-related clinical data from the following:
Limited data sets (LDS) which may include dates (such as admission, discharge, service), date of birth (DOB), date of death (DOD); city, state, five digit or more zip code, and age

De-identified data of medical records that include visits & procedure occurrences, diagnosis, diseases, prescription & drugs, medical devices and supplies, medical conditions, date of death and health care provider locations.

Synthetic data sets that includes an artificial, statistically-comparable, computational derivative of the original data.

These data sets are authorized through Data Transfer Agreements (DTA) & Data Sharing & User Agreements (DUA) titled the NIH CoVid-19 Data Warehouse.

N3C users log in using one of the following:
NIH Login - NIH Identity, Credential, and Access Management Services: Identity Management Services (IMS), which maintains its own unique privacy impact assessment (PIA) on record, including all legal authorities documented.

HHS Personal Identity Verification (PIV), a US Federal government wide credential used to access Federally controlled facilities and information systems at the appropriate security level

Login.gov -a publicly available secure online access to participating government programs

Does the system collect, maintain, use or share PII?

Yes

Indicate the type of PII that the system will collect or maintain.

Date of Birth

Medical Notes

Dates (such as admission, discharge, service)

Date of death (DOD); age

City, state, zip code

Indicate the categories of individuals about whom PII is collected, maintained or shared.

Public Citizens

Patients

How many individuals' PII is in the system?

10,000-49,999

For what primary purpose is the PII used?

Information is for CoVid-related research and geographic impact.

Describe the secondary uses for which the PII will be used.

NA - There is no secondary use.

Identify legal authorities governing information use and disclosure specific to the system and program.

5 USC 301, Departmental regulations.

Are records on the system retrieved by one or more PII data elements?

No

Identify the sources of PII in the system.

Government Sources

Other Federal Entities

Non-Governmental Sources

Public

Private Sector

Identify the OMB information collection approval number and expiration date

Public Law 114-255, Section 2035, exempts research conducted by NIH from Paperwork Reduction Act (PRA) requirements.

Is the PII shared with other organizations?

Yes

Identify with whom the PII is shared or disclosed and for what purpose.

Within HHS

For CoVid 19 research.

Other Federal Agencies

For CoVid 19 research.

Private Sector

The PII will be shared with grantees, the National Center for Data To Health (CD2H) program and other collaborators. They will have access to perform analytical functions on the data collected to help study CoVid related diseases.

Describe any agreements in place that authorizes the information sharing or disclosure.

NCATS is in the process of establishing Data Transfer Agreement (DTA) and Data Sharing and User Agreement partners (data provider and data users) that describes the terms and conditions agreed for sharing and disclosing information. The terms of DTA states that CoVid & CoVid-19 data hosted in N3C is protected from disclosure by a Certificate of Confidentiality per Section 301(d) of the Public Health Service Act and the NIH Policy on Certificates of Confidentiality (CoC).

Describe the procedures for accounting for disclosures.

Not applicable. Data is de-identified or provisioned under the DTA/DUA.

Describe the process in place to notify individuals that their personal information will be collected. If no prior notice is given, explain the reason.

A prior notice is not applicable as NCATS is not involved in the collection process. The DTA prohibits NCATS from attempting to identify or contact individuals who are or may be sources of the LDS without specific written approval from the provider and appropriate Institutional Review Board (IRB).

Is the submission of PII by individuals voluntary or mandatory?

Voluntary

Describe the method for individuals to opt-out of the collection or use of their PII. If there is no option to object to the information collection, provide a reason.

Under the DTA, NCATS is not involved is the collection of the data.

Process to notify and obtain consent from individuals whose PII is in the system when major changes occur to the system.

Under the DTA, NCATS is not involved is the collection of the data.

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.

There is not a process in place to respond to individual concerns if they believe the PII had been inappropriately obtained, used, disclosed or inaccurate. The DTA identified the provider as the legal authority to collect and share the data with the recipient.

Describe the process in place for periodic reviews of PII contained in the system to ensure the data's integrity, availability, accuracy and relevancy.

There is not a defined process in place for periodic review of PII data in N3C. As specified in the DTA 'Terms and Conditions', the information is understood to be provided is 'AS IS'. The DUA stipulates under section 'The Data contributor retains ownership of the Data' that the data is made accessible "AS IS' and NCATS and the data contributors make no representation or warranties.

Identify who will have access to the PII in the system and the reason why they require access.

Users:

For scientific research

Administrators:

To provide support for the system

Contractors:

Direct contractors support and manage the system

Describe the procedures in place to determine which system users (administrators, developers, contractors, etc.) may access PII.

Access to PII is assigned to personnel based upon current job responsibilities. The system uses NIH IMS login to assign permissions/user roles which is considered Personally Identifiable Information (PII).

Describe the methods in place to allow those with access to PII to only access the minimum amount of information necessary to perform their job.

Determinations are made based on role-based access controls and least privilege. User rights are provisioned based on controls within the system, allowing users only access to the minimum amount of PII necessary to perform their job.

Identify training and awareness provided to personnel (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.

The NIH Security Awareness Training course is used to satisfy this requirement. According to NIH policy, all personnel who use NIH applications must attend security awareness training every year. There are five categories of mandatory IT training (Information Security, Counterintelligence, Privacy Awareness, Records Management and Emergency Preparedness).

Describe training system users receive (above and beyond general security and privacy awareness training).

Users are provided in-platform walk-through training that covers core functionality and appropriate and authorized workflows in the N3C system. All data used and displayed in training is hypothetical.

Do contracts include Federal Acquisition Regulation and other appropriate clauses ensuring adherence to privacy provisions and practices?

Yes

Describe the process and guidelines in place with regard to the retention and destruction of PII.

LDS of Patient Medical Records-DAA-0443-2012-0007-0010 (The cut of us annually after 5 years of inactivity ad destroy when case is no longer needed for scientific reference.

I-0012: Pathology Test Records - DAA-0443-2012-0007-0012 (The cut off annually after the date of reporting. Destroy 10 years after cutoff)

Imaging records - DAA-0443-2012-0007-0007
(Cut off in 5 year intervals by fiscal year after file becomes inactive or when no longer needed for clinical reference, whichever is longer. Destroy 60 years after cutoff)

Describe, briefly but with specificity, how the PII will be secured in the system using administrative, technical, and physical controls.

Administrative
Users log in using secure government portals. Platform and data access are determined by the appropriate administrative authorities as described in the N3C Internal Review Board (IRB) protocol. Administrative controls are reinforced by the technical and physical controls laid out below.

Technical
Palantir Foundry Cloud Service (PFC), a Fed-ramp authorized cloud service & contractor's management service of N3C, ensures that data is secured in the system via several technical means. Across the platform, data is encrypted in transit and at rest. Palantir provides highly configurable access controls. Palantir Foundry platform supports comprehensive auditing of all data processing and access. It captures metadata about the source of all data and maintains records of data imports, reads, writes, searches, exports, and deletions.

Physical
PFC is a cloud service platform hosted in Amazon Web Services (AWS Govcloud). PFC system architecture is aligned with a number of framework and classified as a moderate security category with a FedRAMP Authorization to Operate and host a sensitive data.

Identify the publicly-available URL:

unite.nih.gov

Note: web address is a hyperlink.

Does the website have a posted privacy notice?

Yes

Is the privacy policy available in a machine-readable format?

Yes

Does the website use web measurement and customization technology?

Yes

Select the type of website measurement and customization technologies is in use and if it is used to collect PII.

Session Cookies that do not collect PII.

Persistent Cookies that do not collect PII.

Does the website have any information or pages directed at children under the age of thirteen?

No

Does the website contain links to non- federal government websites external to HHS?

No

Is a disclaimer notice provided to users that follow external links to websites not owned or operated by HHS?

null