US Department of Health and Human Services

Privacy Impact Assessment

Date Signed:

08/08/2024

OPDIV:

NIH

Name:

NCATS NCATS Integrated Data Analysis Portal (NIDAP)

PIA Unique Identifier:

P-5878036-950930

The subject of this PIA is which of the following?

Minor Application (stand-alone)

Identify the Enterprise Performance Lifecycle Phase of the system.

Operations and Maintenance

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?

No

Identify the operator.

Agency

Is this a new or existing system?

New

Does the system have Security Authorization (SA)?

Yes

Indicate the following reason(s) for updating this PIA.

Describe the purpose of the system.

NIH Integrated Data Analysis Platform (NIDAP) is a cloud-based, collaborative data aggregation and analysis platform that hosts user-friendly bioinformatics workflows and component analysis and visualization tools. NIDAP was developed by the National Center for Advancing Translational Sciences (NCATS) and the National Cancer Institute (NCI) to provide advanced data management functionality and support, in-network sharing by enabling controlled access to pre-submission data sharing, and integration of open-source software that allows researchers to collaborate and perform basic bioinformatics analysis.

As a scientific resource, NIDAP brings together data, storage, compute, and broadly useful, customizable, data analysis tools in a centralized location available to all NCATS intramural investigators and their external collaborator(s).

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

NIDAP collects business contact information in the form of a name, email address, and roles/title using NCATS Unified Authentication (NCATS UNA). NCATS UNA is an identity brokering hub service provider that establishes trust relationships with other identity providers such as NIH Identity, Credential and Access Management (IAM) Services, HHS Personal Identity Verification (PIV), Login.gov, and partnering Universities and research associates.

In addition to a business contact, NIDAP receives and maintains personally identifiable information (PII) from the following clinical information systems: NCI HALO, Cancer.gov and the NIH Biomedical Translational Research Information System (BTRIS). Information can include name, Medical Record Number (MRN), date of birth (DOB), medical records/notes and genomics sequencing data. Referring physician information is also included in some cases and may consist of name, email, physical addresses, and phone number(s).

Drug related information maintained, stored and/or shared, include compound screening data from NCATS projects, compound structural data and targeted drug information from external sources such as ChEMBL and DrugBank, NCATS project data and data directly from ClinicalTrials.gov.

BTRIS, NCATS UNA, Clinicaltrials.gov, NIH IAM Services, HHS PIV and Login.gov maintain their own unique privacy impact assessment (PIA) on record, with all legal authorities documented. ChEMBL is a manually curated chemical database of bioactive molecules with drug inducing properties. It is maintained by the European Bioinformatics Institute (EBI), of the European Molecular Biology Laboratory (EMBL), based at the Wellcome Trust Genome Campus, Hinxton, UK. HALO is a product (not an acronym) of Indica Labs in collaboration with NCI.

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

NIDAP is a cloud-based, collaborative data aggregation and analysis platform that hosts user-friendly bioinformatics workflows and component analysis and visualization tools. NIDAP was developed by NCATS and NCI to provide advanced data management functionality and support, in-network sharing by enabling controlled access to pre-submission data sharing, and integration of open-source software that allows researchers to collaborate and perform basic bioinformatics analysis. As a scientific resource, NIDAP brings together data, storage, compute, and broadly useful, customizable, data analysis tools in a centralized location available to all NCATS intramural investigators and their external collaborator(s).

NIDAP collects business contact information in the form of a name, email address, and roles/title using NCATS UNA. NCATS UNA is an identity brokering hub service provider that establishes trust relationships with other identity providers such as NIH IAM Services, HHS PIV, Login.gov, and partnering Universities and research associates.

In addition to a business contact, NIDAP receives and maintains PII from the following clinical information systems: NCI HALO, Cancer.gov and the NIH BTRIS. Information can include name, MRN, DOB, medical records/notes and genomics sequencing data. Referring physician information is also included in some cases and may consist of name, email, physical addresses, and phone number(s).

Drug related information maintained, stored and/or shared, include compound screening data from NCATS projects, compound structural data and targeted drug information from external sources such as ChEMBL and DrugBank, NCATS project data and data directly from ClinicalTrials.gov.

BTRIS, NCATS UNA, Clinicaltrials.gov, NIH IAM Services, HHS PIV and Login.gov maintain their

own unique privacy impact assessment (PIA) on record, with all legal authorities documented. ChEMBL is a manually curated chemical database of bioactive molecules with drug inducing properties. It is maintained by the EBI, of the EMBL, based at the Wellcome Trust Genome Campus, Hinxton, UK. HALO is a product (not an acronym) of Indica Labs in collaboration with NCI.

NIDAP is hosted in Palantir Foundry Cloud Service (PFCS), a Federal Risk and Authorization Management Program (FedRAMP) authorized data-driven integration and analytical Software as- a-Service (SaaS) platform.

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
Name
E-Mail Address
Mailing Address
Phone Numbers
Medical Records Number
Medical Notes
Role/Title
Genomics sequencing data, compound screening data

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

Employees
Public Citizens
Business Partner/Contacts (Federal/state/local agencies)

How many individuals' PII is in the system?

50,000-99,999

For what primary purpose is the PII used?

PII is used to grant access to specific resources in NIDAP and for research.

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

NA

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

Public Health Service Act. (42 U.S.C. 241, 242, 248, 281, 282, 284, 285a-285q, 287, 287b, 287c, 289a, 289c, and 44 U.S.C. 3101)

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

Yes

Identify the number and title of the Privacy Act System of Records Notice (SORN) that is being use to cover the system or identify if a SORN is being developed.

09-25-0099 Clinical Research: Patient Medical Records

Identify the sources of PII in the system.

Directly from an individual about whom the information pertains

Online

Government Sources

Within the DOC

Non-Government 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?

No

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

NCATS UNA provides a consent statement, along with a privacy statement before a user has access. (https://nidap.nih.gov/multipass/login/all)

Patient/clinical information is pulled from the various source systems and is the responsibility of those systems.

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.

If an individual wants access to NIDAP resources, they would have consented with their identity provider to share the requested information. Otherwise, they don't have access.

Patient/clinical information is pulled from the various source systems and is the responsibility of those systems.

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

Changes to the policy would be posted on the privacy page or at the time of login.

Patient/clinical information is pulled from the various source systems and is the responsibility of those systems.

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.

A central email address, NCATSAuthSupport@mail.nih.gov, is available to resolve users' concerns regarding PII. NIDAP-support@nih.gov is available for users wanting direct NIDAP support.

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

NCATS will conduct an annual review of information in the system to ensure integrity, accuracy and relevancy of the data.

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

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 job responsibilities and a need-to-know basis. An NCATS UNA account is required to gain access to the stored PII data.

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.

According to NIH policy, all personnel who use NIH applications must attend security awareness training every year. There are five categories of mandatory information technology (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).

NCATS direct contractors have completed the NIH Role-based training for IT System Administrators.

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

Records are retained and disposed of under the authority of the NIH Records Retention Schedule.

Item 07-203 - System access records; Systems not requiring special accountability for access.
These records are created as part of the user identification and authorization process to gain access to systems. Records are used to monitor inappropriate systems access by users.
Exclusion 1. Excludes records relating to electronic signatures.
Exclusion 2. Does not include monitoring for agency mission activities such as law enforcement.
Disposition: Destroy when business use ceases. DAA-GRS-2013-0006-0003

Item 07-204 - System access records. Systems requiring special accountability for access.
These records are created as part of the user identification and authorization process to gain access to systems. Records are used to monitor inappropriate systems access by users.
Exclusion 1. Excludes records relating to electronic signatures.
Exclusion 2. Does not include monitoring for agency mission activities such as law enforcement.
Systems requiring special accountability for access.
These are user identification records associated with systems which are highly sensitive and

potentially vulnerable.
Disposition: Destroy 6 years after password is altered or user account is terminated, but longer retention is authorized if required for business use. DAA-GRS-2013-0006-0004

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

Administrative Controls:
Users log in using a secure government portal. Platform and data access are determined by the appropriate administrative authorities.

Technical Controls:
Palantir Foundry Cloud Service, a FedRAMP authorized cloud service & contractor's management service for NIDAP ensures that data is secured in the system via several technical means. Across the platform, data is encrypted in transit and at rest. Palantir provided highly configurable success controls. The NIDAP Palantir Foundry platform support comprehensive auditing of all data process and maintains record of data imports, reads, writes, searches, exports and deletions. Palantir Foundry is aligned with a number of framework and classified as a moderate security category with FedRAMP Authorization to Operate and host sensitive data.

Physical Controls: NIDAP is a Palantir cloud service platform hosted in Amazon Web Services US East.