US Department of Health and Human Services

Privacy Impact Assessment

Date Signed:

10/26/2016

OPDIV:

FDA

Name:

CDRH High Performance Computing

PIA Unique Identifier:

P-4718783-381578

The subject of this PIA is which of the following?

General Support System (GSS)

Identify the Enterprise Performance Lifecycle Phase of the system.

Operations and Maintenance

Is this a FISMA-Reportable system?

Yes

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.

The High Performance Computing (HPC) system is an information technology investment that is operated and maintained by the FDA's Center for Devices and Radiological Health (CDRH) Office of Scientific and Engineering Laboratories (OSEL). It provides high-performance computational clusters specifically engineered to facilitate large-scale modeling and simulations needed by FDA scientists to protect and promote public health. The HPC provides capabilities for regulatory science and also for regulatory activities.

The clusters enable analytic programs to run at a speed not obtainable with a scientific workstation and at a scale not previously possible. HPC also provides a secure environment for the intermediate work products of research and development in FDA regulatory science. Additionally, it will store working data and intermediate work products of genomic analyses performed in support of certain regulatory activities.

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

HPC provides secure storage for a wide variety of data required in the development of regulatory methods, procedures and standards related to FDA's public health mission. Some projects performed with the HPC may require bioinformatics data, including electronic patient records and partial or complete genome files. Most significantly, HPC is engineered to receive and analyze exceptionally large genetic sequencing datasets. Patient information for certain protocols may include: the demographics of the patient such as age (possibly date of birth), sex, and ethnicity; health conditions, status, or treatment of an individual; and source of the data (possibly including specific facilities, which may provide some insight into geographic location of the subject). For all projects, analysts are directed to use the minimum amount of PII needed to conduct the specific study, and in particular not to use or maintain names, contact information, or unique identifying numbers that can be traced to an individual, unless such identifiers are needed for the objectives of a specific study. Social Security numbers are not needed and are not one of the data elements used.

Some users will access the HPC using a system-specific username and password, and these access credentials will be contained in the system. Others will be able to access the system using a different single sign-on multi-factor authentication approach. Users are all FDA employees and not direct or non-direct contractors.

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

The HPC consists of two clusters on which FDA scientists perform resource-intensive computational projects such as molecular and fluid dynamics simulations, Bayesian statistical analyses, and genomic analyses. The work is completed with unusually high speeds because the HPC contains an internal management function which "parallelizes" the project elements and runs them on a large number of compute nodes simultaneously.

The HPC stores various reference data sets needed by the modeling and simulation software programs used by FDA scientists. In most cases, these data sets relate to basic science such as microbiology, chemistry, and physics. The HPC also stores the intermediate work products generated by research projects.

For certain scientific investigations (i.e., detailed analysis), the HPC may collect PHI and PII. Due to existing regulatory restrictions (i.e., HIPAA, Privacy Act, and others) on the use of data that contains PHI and PII, the FDA will request only the minimum necessary to achieve its mission, in agreement with the partners from which the data is obtained.

The source of the data provided may be public health agencies; research institutions; research universities; hospitals; clinical professional societies; clinical research organizations; and other similar private organizations. None of this information is sent directly to HPC, however, so no interconnection security agreements (ISAs) concerning this system and systems outside of the FDA is needed.

The HPC environment also stores user session information in a protected area. User access is controlled by multi-factor authentication (including use of Personal Identity Verification (PIV) card). To grant a new user an HPC account, the HPC system administrator assigns a unique HPC user ID. The user must enter this user ID plus a personal secret password to begin an HPC session. The results of the HPC login are logged for later tracking/reporting of user access to the environment. The logging includes failed login attempts.

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 Records Number

Medical Notes

Device Identifiers

Specific identifiers collected vary based on the relevant medical device and data source.

User credentials

Device identifiers refers to medical device identifiers.

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

Employees

Patients

How many individuals' PII is in the system?

1,000,000 or more

For what primary purpose is the PII used?

PII is used in the context of epidemiological and analytical research (study and analysis of the patterns, causes, and effects of health and disease conditions in defined populations) to allow grouping, sorting and manipulation of metadata in the post-market monitoring of medical products to generate statistics and identify broad trends. Usernames and passwords are used for authentication.

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

Not applicable.

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

The Federal Food, Drug and Cosmetics Act, 21 U.S.C. 301 et seq. This project allows FDA/HPC to protect and promote public health. See also 45 CFR 164.512 , the Public Health Exception to the HIPAA Rule. This authorizes FDA to collect patient medical records.

In addition, the security and privacy measures of the system, including the use of usernames and passwords, are required by the Federal Information Security Management Act (FISMA) and the statutes underlying OMB Circular A-130 for the secure and efficient use of government systems and resources.

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

No

Identify the sources of PII in the system.

Government Sources

Within OpDiv

Other HHS OpDiv

State/Local/Tribal

Foreign

Non-Governmental Sources

Private Sector

Identify the OMB information collection approval number and expiration date

Not applicable.

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.

All personal information in the HPC environment arrives under existing statutory authorities from regulated industry or private registries who will be referred to hereinafter as the "data sources." Informed consent responsibilities rest with the parties that have collected the data and FDA has no way to contact the data subjects directly.

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.

All personal information in the HPC environment arrives, under existing statutory authorities, from regulated industry or private registries who will be referred to hereinafter as the "data sources." Informed consent responsibilities and the provision of any opt-out methods rest with the parties that have collected the data and FDA has no way to contact the data subjects directly.

System users are employees and the system contains PII relevant to their access credentials. There is no method for employees to opt not to submit PII. Permanent employees, direct contract employees, fellows and other personnel must provide their PII in order for the Agency to process administrative materials and securely administer access to Agency information and property.

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

HPC will follow agreed-upon and approved procedures to notify the data sources of any relevant changes. The data sources in regulated industries will then notify the individuals as necessary. FDA is obligated to abide by the minimum necessary standard for request, acquisition, and use of PII. All requests and transfers of PII to FDA systems must adhere to a pre-specified analysis protocol that documents the minimum necessary PII required to enable analyses that address specific public health questions. These questions are formulated on a population basis such that PII is used to identify groups for analysis, but never to access individual records. Results and reports of the analyses never contain PII. Almost all information in this system is de-identified and could not in any circumstances be used to locate or contact any specific individual.

If the Agency ever makes other use of the PII of its employees that serve as access credentials, employees could be contacted many ways, including by phone, e-mail, and web notices.

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.

HPC is not collecting PII directly from individuals. HPC interaction is restricted to the information provided by the data sources.

Individuals who suspect their PII has been inappropriately obtained, used or disclosed in any FDA system have a number of avenues available to request to rectify the situation. Often, these individuals contact the office or division where they have determined that their information is held. Individuals may then make further requests for their information to be corrected or amended. FDA considers these requests and, if appropriate, makes the requested changes.

Employees with such concerns can additionally work with their supervisors, a 24-hour technical assistance line, FDA's Systems Management Center, and other channels.

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

Each project that uses PII is reviewed annually. The review addresses the continued relevancy of any PII being used as well as data integrity and accuracy. The FDA Office of Information Management and Technology (OIMT) conducts separate security reassessment of the HPC environment annually to audit the technology, policies and operational controls implemented by the HPC. The OIMT assessors designate and review a varying subset of high-priority controls of data integrity, availability, accuracy, and relevancy. The results of their audit serve as the basis for renewing the HPC's authority to operate (ATO). In daily operations, the project's technology controls and approved procedures protect data integrity in all data access operations and data transfers.

For logon credentials, FDA personnel are responsible for providing accurate information and may independently update and correct their information at any time.

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

Users:

Users may have access to any PII contained in research data, if any, as described previously. Users have access to their own access credentials.

Administrators:

May perform random spot checks for integrity verification upon user request. May have access to others' usernames but not passwords.

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

HPC employs Role Based Access Control (RBAC) and data set segmentation to control access to data. Access to the segment of data for a given project must be granted by explicit authorization of the Principal Investigator (PI, the lead researcher) in charge of the project.

Administrators may have access to usernames in the course of creating and maintaining accounts. They will not have access to passwords.

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.

HPC employs RBACs and data set segmentation to ensure that the user can access data only with a valid need-to-know, and a given data segment is accessible only by authorized users.

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.

All HPC users must go through the FDA-wide annual Security Awareness Training. FDA-wide training covers rules of behavior (ROB) and mandatory FDA rules for protection of data confidentiality.

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

HPC requires its users to undergo a one (1) hour minimum training session which includes understanding of and compliance with required methods of protecting data confidentiality as specifically applicable to the HPC environment. Additionally, HPC system administrators receive yearly refresher training in the HPC system administrator rules of behavior and current best practices in IT system security, including the protection of data confidentiality.

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.

Data retention schedules will vary by project. Individual investigators are required to name a Data Manager. The Data Manager is responsible for working with an Assistant Records Liaison Officer (ARLO) to identify an appropriate records schedule, or to create one and receive National Archives and Records Administration (NARA) approval. Data Managers must also ensure that retention schedules are being implemented and followed. As a general guideline, however, HPC will use FDA’ s File Code 4230, Research Project Working Files (NARA approval N1 088-04-5). Under this schedule, disposition is temporary. Cut off date for these records is at end of the calendar year in which the project is completed, after which files are maintained for three years, then destroyed or deleted six years after the cutoff date, or when no longer needed for reference, whichever is sooner.

Logon credentials remain available as long as each user has authorized access to the system. Credentials are revoked immediately when access is no longer needed, including if the individual leaves FDA employment.

Retention is maintained under FDA File Code 9962 (NARA GRS 20, Item 1c; superseded by the new GRS 3.2, item 030 (DAA-GRS-2013-0006-0003), which is for “records … 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. This schedule applies to records such as “user profiles; log-in files; password files; audit trail files and extracts; system usage files; and cost-back files used to assess charges for system use.”

Under this schedule, retention is until "when business use ceases." In other words, NARA concurs that agencies may dispose of these records as soon as they are no longer needed.

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

Administrative controls include role-based access controls (RBAC), annual training, and periodic reviews of access control lists. Physical controls include that all servers are contained in locked offices within a secure facility with full-time guards and closed circuit cameras.

Specific controls in place to protect PII are as follows:

User credentialing and authorization- Users are vetted through an FDA approval process (i.e., background check, including finger print) prior to accessing the system. HPC employs RBAC, separation of duties, and least privilege.

HPC uses an access control list (ACL). This list contains the data types typically tied to a project, the users designated to access the data, and user permission level (e.g., read only, read/write, etc.) This list is reviewed by senior management at least annually, and when there is a significant change (e.g., change in staff).

Group accounts or account sharing are not allowed.

Multi-factor authentication and access control. All access to the HPC requires multi-factor authentication and is PIV-controlled. The use of the PIV card allows for the mapping of users and their activities. For researchers who use applications reside within the HPC's native environment, they must authenticate with their unique HPC user ID and password. Their login invokes a layer of protection, tracking and accountability during their entire login session.

Defense-In-Depth -- Services that increase the risk of common malware infections are removed from HPC. For example, the HPC environment does not offer e-mail capabilities, nor web browser access to the Internet.

Continuous Monitoring- HPC undergoes an annual risk assessment to ensure that the security controls are implemented and continue to meet requirements.