U.S. Department of Commerce

NOAA

Privacy Threshold Analysis
for the
NOAA Research & Development High Performance Computing System (R&D HPCS) – NOAA0500

U.S. Department of Commerce Privacy Threshold Analysis

NOAA Research & Development High Performance Computing System

(R&D HPCS)

Unique Project Identifier: 006-000380400 00-48-01-17-01-00

Introduction: This Privacy Threshold Analysis (PTA) is a questionnaire to assist with determining if a Privacy Impact Assessment (PIA) is necessary for this IT system. This PTA is primarily based from the Office of Management and Budget (OMB) privacy guidance and the Department of Commerce (DOC) IT security/privacy policy. If questions arise or further guidance is needed in order to complete this PTA, please contact your Bureau Chief Privacy Officer (BCPO).

Description of the information system and its purpose: NOAA Research & Development High Performance Computing System (R&D HPCS), NOAA0500 provides research and development weather models in support of NOAA's operational mission. The R&D HPCS operates large scale, extreme computing environments that encompass multiple geographic sites, and heterogeneous supercomputing architectures. This system supports NOAA's mission by providing cutting edge technology for weather and climate model developers. These models eventually form the basis for NOAA's daily weather forecasts, storm warnings, and climate change forecasts. System users include scientists from multiple NOAA Line Offices, and their research collaborators, including some foreign nationals. NOAA's R&D HPC system (R&D HPCS) provides four fundamental HPC functions:

1. Large-scale computing provides computing for development, testing, and production integrations of NOAA environmental models. The workload that runs on this subsystem is characterized by compute-intensive codes with I/O characterized by regular snapshots of diagnostic fields.

2. Analysis and interactive computing provides computing for the post-processing of data from production runs and the analysis of post-processed data, code development, and debugging. The workload that runs on this subsystem is characterized by data-intensive codes requiring high I/O bandwidth.

3. Data archiving provides long-term storage of post-processed model runs and analyses.

4. Networking links these subsystems together.

The E-Government Act of 2002 defines "information system" by reference to the definition section of Title 44 of the United States Code. The following is a summary of the definition: "Information system" means a discrete set of information resources organized for the collection, processing, maintenance, use, sharing, dissemination, or disposition of information. See: 44. U.S.C. § 3502(8).

a) *Whether it is a general support system, major application, or other type of system*

NOAA0500 system is a General Support System (GSS).

b) *System location*

The current configuration of the R&D HPCS is architected along organizational lines. Large-scale computing, analysis computing, and storage are located at the following locations:
1. NOAA Earth System Research Laboratory (ESRL), David Skaggs Research Center, Boulder, CO
2. NOAA Geophysical Fluid Dynamics Laboratory (GFDL), Princeton University Forrestal Campus, Princeton, NJ
3. NOAA Environmental, Security Computing Center (NESCC), Fairmont, WV

c) *Whether it is a standalone system or interconnects with other systems (identifying and describing any other systems to which it interconnects)*

Interconnections with Oak Ridge National Laboratories (ORNL) and NOAA8866 and documented via ISA.

d) *The purpose that the system is designed to serve*

The purpose of NOAA's R&D HPC system (R&D HPCS) is to provide four fundamental HPC functions:
1. Large-scale computing provides computing for development, testing, and production integrations of NOAA environmental models. The workload that runs on this subsystem is characterized by compute-intensive codes with I/O characterized by regular snapshots of diagnostic fields.
2. Analysis and interactive computing provides computing for the post-processing of data from production runs and the analysis of post-processed data, code development, and debugging. The workload that runs on this subsystem is characterized by data-intensive codes requiring high I/O bandwidth.
3. Data archiving provides long-term storage of post-processed model runs and analyses.
4. Networking links these subsystems together.

e) The way the system operates to achieve the purpose

These users access the system and submit weather or climate modeling application program runs via job scheduling software. These models contain parallelized code to take advantage of the large scale, highly parallelized environment offered by the HPCS. This is necessary to support the science. Modeling jobs can be extremely large (e.g., 1200 processors required), because they incorporate different local, regional, or global atmospheric and ocean models to create an ensemble model program. The scheduling software automatically identifies and collects the necessary processors to run the job, and controls its execution. Therefore, the user never has any direct interaction with the compute nodes of the system.

Weather and climate data is collected from a variety of sources and fed into the system by the user community. All of this data is vetted by the NOAA scientific community through processes outside of this system, to guarantee its authenticity and integrity. Once entered into the R&D HPCS, the system security controls are designed to guarantee the integrity of this data.

f) A general description of the type of information collected, maintained, use, or disseminated by the system

The National Centers for Environmental Prediction (NCEP) utilize data acquired from commercial, other U.S Government and International sources to execute NCEP mission. A subset of this data, referred to as "restricted data" is made available to NCEP with restrictions on further dissemination.* As a direct or indirect party to the agreements governing the use of this Restricted Data, NCEP is charged with protecting restricted data during use and identifying restricted data to managers, users, staff and partners supporting NCEP mission. Authority for collection of information: 5 U.S.C. 301 5 U.S.C. 301 authorizes the operations of an executive agency, including the creation, custodianship, maintenance and distribution of records.

*NOAA has agreements with ships and planes, which collect local weather data while at sea/in the air and share with NOAA. The data includes the positions of those ships and planes, because the two types of information cannot be separated. The location data is considered proprietary.

Information sharing: The PII in the system will not be shared outside of the bureau except in case of a breach. The BII (restricted data). NCEP receives and shares with RDHPCS.

Account-related information, consisting of First and Last Name, affiliation, phone, email address, organization and organizational mailing address and if a Foreign National, we also collect Country of Citizenship, Country of Current Residence and Home Country, is stored in a protected database with limited access. User may update their information at any time, but retrieving the entire database is limited. It can be accessed via secure web browsing sessions. Account-related information is used in the creation/management of NOAA0500 user accounts.

g) *Identify individuals who have access to information on the system*

Restricted Production (RSTPROD) data is only accessible by a set of exclusively authorized individuals consisting of Federal employees and contractors working on behalf of NOAA0500 and NOAA.

Account-related information is accessible by the user and administrators of the system.

h) *How information in the system is retrieved by the user*

RSTPROD data is accessed via a secure channel (SecureShell tunnel) and via membership to the project.

Account-related information is retrieved by the user via a web browser and accessing the Account Information Management (AIM) database.

i) *How information is transmitted to and from the system.*

Securely via the use of SecureShell (ssh) tunnels for RSTPROD.

Securely via the use of SSL/ TLS for AIM (account-related) content.

Questionnaire:

1. What is the status of this information system?

____ This is a new information system. *Continue to answer questions and complete certification.*

____ This is an existing information system with changes that create new privacy risks. *Complete chart below, continue to answer questions, and complete certification.*

Changes That Create New Privacy Risks (CTCNPR)					
a. Conversions		d. Significant Merging		g. New Interagency Uses	
b. Anonymous to Non-Anonymous		e. New Public Access		h. Internal Flow or Collection	
c. Significant System Management Changes		f. Commercial Sources		i. Alteration in Character of Data	
j. Other changes that create new privacy risks (specify):					

____ This is an existing information system in which changes do not create new privacy risks, and there is not a SAOP approved Privacy Impact Assessment. *Continue to answer questions and complete certification.*

____ This is an existing information system in which changes do not create new privacy risks, and there is a SAOP approved Privacy Impact Assessment (version 01-2015). *Continue to answer questions and complete certification.*

_X_ This is an existing information system in which changes do not create new privacy risks, and there is a SAOP approved Privacy Impact Assessment (version 01-2017 or later). *Skip questions and complete certification.*

2. Is the IT system or its information used to support any activity which may raise privacy concerns?

NIST Special Publication 800-53 Revision 4, Appendix J, states "Organizations may also engage in activities that do not involve the collection and use of PII, but may nevertheless raise privacy concerns and associated risk. The privacy controls are equally applicable to those activities and can be used to analyze the privacy risk and mitigate such risk when necessary." Examples include, but are not limited to, audio recordings, video surveillance, building entry readers, and electronic purchase transactions.

____ Yes. *(Check all that apply.)*

Activities			
Audio recordings		Building entry readers	
Video surveillance		Electronic purchase transactions	
Other (specify):			

_X_ No.

3. Does the IT system collect, maintain, or disseminate business identifiable information (BII)?

As per DOC Privacy Policy: "For the purpose of this policy, business identifiable information consists of (a) information that is defined in the Freedom of Information Act (FOIA) as "trade secrets and commercial or financial information obtained from a person [that is] privileged or confidential." (5 U.S.C.552(b)(4)). This information is exempt from automatic release under the (b)(4) FOIA exemption. "Commercial" is not confined to records that reveal basic commercial operations" but includes any records [or information] in which the submitter has a commercial interest" and can include information submitted by a nonprofit entity, or (b) commercial or other information that, although it may not be exempt from release under FOIA, is exempt from disclosure by law (e.g., 13 U.S.C.)."

_X_ Yes, the IT system collects, maintains, or disseminates BII.

____ No, this IT system does not collect any BII.

4. Personally Identifiable Information (PII)

4a. Does the IT system collect, maintain, or disseminate PII?

As per OMB 17-12: "The term PII refers to information that can be used to distinguish or trace an individual's identity either alone or when combined with other information that is linked or linkable to a specific individual."

_X_ Yes, the IT system collects, maintains, or disseminates PII about: *(Check all that apply.)*

_X__ DOC employees
_X__ National Institute of Standards and Technology Associates
_X__ Contractors working on behalf of DOC
_X__ Other Federal Government personnel
_X__ Members of the public

____ No, this IT system does not collect any PII.

If the answer is "yes" to question 4a, please respond to the following questions.

4b. Does the IT system collect, maintain, or disseminate Social Security numbers (SSNs), including truncated form?

____ Yes, the IT system collects, maintains, or disseminates SSNs, including truncated form.

Provide an explanation for the business need requiring the collection of SSNs, including truncated form. n/a

Provide the legal authority which permits the collection of SSNs, including truncated form. n/a

_X__ No, the IT system does not collect, maintain, or disseminate SSNs, including truncated form.

4c. Does the IT system collect, maintain, or disseminate PII other than user ID?

_X__ Yes, the IT system collects, maintains, or disseminates PII other than user ID.

____ No, the user ID is the only PII collected, maintained, or disseminated by the IT system.

4d. Will the purpose for which the PII is collected, stored, used, processed, disclosed, or disseminated (context of use) cause the assignment of a higher PII confidentiality impact level?

Examples of context of use include, but are not limited to, law enforcement investigations, administration of benefits, contagious disease treatments, etc.

____ Yes, the context of use will cause the assignment of a higher PII confidentiality impact level.

_X__ No, the context of use will not cause the assignment of a higher PII confidentiality impact level.

If any of the answers to questions 2, 3, 4b, 4c, and/or 4d are "Yes," a Privacy Impact Assessment (PIA) must be completed for the IT system. This PTA and the approved PIA must be a part of the IT system's Assessment and Authorization Package.

CERTIFICATION

_X___ I certify the criteria implied by one or more of the questions above **apply** to the NOAA Research & Development High Performance Computing System (R&D HPCS) and as a consequence of this applicability, I will perform and document a PIA for this IT system.

_____ I certify the criteria implied by the questions above **do not apply** to the NOAA Research & Development High Performance Computing System (R&D HPCS) and as a consequence of this non-applicability, a PIA for this IT system is not necessary.

Name of Information System Security Officer (ISSO) or System Owner (SO): Justin May

Signature of ISSO or SO: MAY.JUSTIN.NATHANIEL.1039635980 2020.04.08 19:54:53 -06'00' Date: __________

Name of Information Technology Security Officer (ITSO): Charley Obenschain

Signature of ITSO: OBENSCHAIN.CHARLES.THOMAS.1506347293 Digitally signed by OBENSCHAIN.CHARLES.THOMAS.1506347293 Date: 2020.04.22 07:46:41 -04'00' Date: __________

Name of Privacy Act Officer (PAO): Adrienne Thomas

Signature of PAO: THOMAS.ADRIENNE.M.1365859600 Digitally signed by THOMAS.ADRIENNE.M.1365859600 Date: 2020.05.23 18:10:46 -04'00' Date: __________

Name of Authorizing Official (AO): Zachary Goldstein

Signature of AO: GOLDSTEIN.ZACHARY.G.1228698985 Digitally signed by GOLDSTEIN.ZACHARY.G.1228698985 Date: 2020.05.21 17:04:48 -04'00' Date: __________

Name of Bureau Chief Privacy Officer (BCPO): Mark H. Graff

Signature of BCPO: GRAFF.MARK.HYRUM.1514447892 Digitally signed by GRAFF.MARK.HYRUM.1514447892 Date: 2020.06.02 13:43:03 -04'00' Date: __________