The Privacy Office
U.S. Department of Homeland Security
Washington, DC 20528
703-235-0780, pia@dhs.gov
www.dhs.gov/privacy

PRIVACY THRESHOLD ANALYSIS (PTA)

This form is used to determine whether a Privacy Impact Assessment is required.

Please use the attached form to determine whether a Privacy Impact Assessment (PIA) is required under the E-Government Act of 2002 and the Homeland Security Act of 2002.

Please complete this form and send it to your component Privacy Office. If you do not have a component Privacy Office, please send the PTA to the DHS Privacy Office:

Rebecca J. Richards
Director of Privacy Compliance
The Privacy Office
U.S. Department of Homeland Security
Washington, DC 20528
Tel: 703-235-0780

PIA@dhs.gov

Upon receipt, the DHS Privacy Office will review this form. If a PIA is required, the DHS Privacy Office will send you a copy of the Official Privacy Impact Assessment Guide and accompanying Template to complete and return.

A copy of the Guide and Template is available on the DHS Privacy Office website, www.dhs.gov/privacy, on DHSOnline and directly from the DHS Privacy Office via email: pia@dhs.gov, phone: 703-235-0780.

PRIVACY THRESHOLD ANALYSIS (PTA)

Please complete this form and send it to the DHS Privacy Office.
Upon receipt, the DHS Privacy Office will review this form
and may request additional information.

SUMMARY INFORMATION

DATE **submitted for review:** February 3, 2010

NAME **of** Project**: Protected Repository for the Defense of Infrastructure Against Cyber Threats (PREDICT)**

Name of Component: Science and Technology

Name of Project Manager: Douglas Maughan

Email for Project Manager: Douglas.Maughan@dhs.gov

Phone number for Project Manager: 202-254-6145

TYPE **of Project:**

☒ **Information Technology and/or System**[*]

□ **A Notice of Proposed Rule Making or a Final Rule.**

□ **Other: <Please describe the type of project including paper based Privacy Act system of records.>**

[*] The E-Government Act of 2002 defines these terms by reference to the definition sections of Titles 40 and 44 of the United States Code. The following is a summary of those definitions:

- "Information Technology" means any equipment or interconnected system or subsystem of equipment, used in the automatic acquisition, storage, manipulation, management, movement, control, display, switching, interchange, transmission, or reception of data or information. See 40 U.S.C. § 11101(6).
- "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).

Note, for purposes of this form, there is no distinction made between national security systems or technologies/systems managed by contractors. All technologies/systems should be initially reviewed for potential privacy impact.

SPECIFIC QUESTIONS

1. Describe the project and its purpose:

The PREDICT system is a respository of test datasets of Internet traffic data that is made available to approved researchers and managed by an outside contractor serving as the PREDICT Coordination Center. PREDICT was created to help to protect and defend the cyber infrastructure of our country. PREDICT datasets are avaialable to approved researchers who are conducting cyber security research that is in the interests of the United States. PREDICT creates a community of data providers and a structure for disseminating critical data sources to those seeking to field the next generation of cyber defense technologies; PREDICT as the potential to accelerate the R&D community's ability to develop effective and timely cyber defense technologies. It will also enable researchers to more closely monitor the emerging trends and patterns of attacks that propagate across the Internet. Due to PREDICT's inability to manage operations and audit and monitor compliance with PREDICT operational policies and procedures outside the United States, all research and work invloving PREDICT datasets must be carried out at locations within the 50 United States.

PREDICT's goals are to

- Provide a central repoistory, accessible through a Web-based portal that catalogs current computer network operational data. (Data Catalog/Metadata including domain name server root server data, Internet topology measurement data, blackhole address space data, Border Gateway Protocol (BGP) upate message and BGP routing table dumps, packet headers, worm data, and firewall logs). Most of the data sets do not contain PII, though some may have IP addresses.

- Provide secure access to multiple sources of data collected as a result of use and traffic on the Internet.

- Facilitate data flow among PREDICT participants for the purpose of developing new models, technolgies and products that support effective threat assessment and increase cyber security capabilities.

2. Status of Project:

☐ This is a new development effort.

☒ This is an existing project.

Date first developed: January 1, 2007

Date last updated:

<Please provide a general description of the update.>

3. **Could the project relate in any way to an individual?**[1]

□ No. Please skip ahead to the next question.

☒ Yes. Please provide a general description, below.

PREDICT requires for you to submit personal information to obtain access to the system and be given an account. The personal information of the individual is stored on the Structured Query Language database. Individuals requesting an account must submit a Sponsorship Letter with this account request. Individuals requesting accounts must select a role to associate the user with specified access to specified data. All roles will need to complete a Memorandum of Agreement to fully participate in PREDICT. The following are the different types of roles associated with PREDICT:

Data Hosts: An organization providing computing infrastructure to host PREDICT datasets

Data Providers: An organization providing datasets to PREDICT

Researchers: Individuals seeking access to PREDICT datasets to develop products or services that support strong cyber defense.

4. **Do you collect, process, or retain information on: (Please check all that apply)**

□ DHS Employees

□ Contractors working on behalf of DHS

☒ The Public

□ The System does not contain any such information.

[1] Projects can relate to individuals in a number of ways. For example, a project may include a camera for the purpose of watching a physical location. Individuals may walk past the camera and images of those individuals may be recorded. Projects could also relate to individuals in more subtle ways. For example, a project that is focused on detecting radioactivity levels may be sensitive enough to detect whether an individual received chemotherapy.

5. Do you use or collect Social Security Numbers (SSNs)? (This includes truncated SSNs)

☒ No.

□ Yes. Why does the program collect SSNs? Provide the function of the SSN and the legal authority to do so:

<Please provide the function of the SSN and the legal authority to do so.>

6. What information about individuals could be collected, generated or retained?

USER Information

Required fields are marked with *

*First Name:

*Last Name:

*Street 1:

Street 2:

*City:

*State:

*Zip Code:

*Phone Number(s):

Office

Home

Cell

Fax:

*E-mail:

SPONSORING ORGANIZATION INFORMATION

Sponsoring Organization

*Organization Name:

*Street 1:

Street 2:

*City:

*State:

*Zip Code:

Authorized Representative

*First Name:

*Last Name:

*Phone:

*E-mail:

7. **If this project is a technology/system, does it relate solely to infrastructure? [For example, is the system a Local Area Network (LAN) or Wide Area Network (WAN)]?**

☒ No. Please continue to the next question.

□ Yes. Is there a log kept of communication traffic?

□ No. Please continue to the next question.

□ Yes. What type of data is recorded in the log? (Please choose all that apply.)

□ Header

□ Payload Please describe the data that is logged.

<Please list the data elements in the log.>

8. **Can the system be accessed remotely?**

□ No.

☒ Yes. When remote access is allowed, is the access accomplished by a virtual private network (VPN)?

□ No.

☒ Yes.

9. Is Personally Identifiable Information[2] physically transported outside of the LAN? (This can include mobile devices, flash drives, laptops, etc.)

☒ No.

□ Yes.

10. Does the system connect, receive, or share Personally Identifiable Information with any other DHS systems[3]?

☒ No

□ Yes. Please list:

11. Are there regular (ie. periodic, recurring, etc.) data extractions from the system?

☒ No.

□ Yes. Are these extractions included as part of the Certification and Accreditation[4]?

- □ Yes.
- □ No.

12. Is there a Certification & Accreditation record within OCIO's FISMA tracking system?

□ Unknown.

□ No.

☒ Yes. Please indicate the determinations for each of the following:

Confidentiality:	□ Low	☒ Moderate	□ High	□ Undefined
Integrity:	□ Low	☒ Moderate	□ High	□ Undefined
Availability:	☒ Low	□ Moderate	□ High	□ Undefined

[2] Personally Identifiable Information is information that can identify a person. This includes; name, address, phone number, social security number, as well as health information or a physical description.

[3] PII may be shared, received, or connected to other DHS systems directly, automatically, or by manual processes. Often, these systems are listed as "interconnected systems" in TAFISMA.

[4] This could include the Standard Operation Procedures (SOP) or a Memorandum of Understanding (MOU)

PRIVACY THRESHOLD REVIEW

(To be Completed by the DHS Privacy Office)

DATE **reviewed by the DHS Privacy Office: April 7, 2010**

NAME **of** the DHS Privacy Office Reviewer: **Rebecca J. Richards**

DESIGNATION

☐ **This is NOT a Privacy Sensitive System** – the system contains no Personally Identifiable Information.

☒ **This IS a Privacy Sensitive System**

Category of System

☒ IT System

☐ National Security System

☐ Legacy System

☐ HR System

☐ Rule

☐ Other:

Determination

☐ PTA sufficient at this time

☐ Privacy compliance documentation determination in progress

☐ PIA is not required at this time

☒ A PIA is required

☒ System covered by existing PIA: S&T Predict

☐ A new PIA is required.

☐ A PIA Update is required.

☒ A SORN is required

☒ System covered by existing SORN: DHS/ALL-002

☐ A new SORN is required.

DHS PRIVACY OFFICE COMMENTS