

**SUPPORTING STATEMENT FOR**  
**PREDICT System Information Collection**  
**(OMB No. 1640-0012)**

**A. Justification**

1. The PREDICT initiative of the Department of Homeland Security (DHS) Science and Technology (S&T) directorate facilitates cyber defense research and development through the establishment of distributed repositories of security-relevant computer and network operations data and making such data available for use by researchers. The PREDICT Coordinating Center (PCC) has established application procedures, protection policies, and review processes necessary to make this data available to the cyber defense research community. PREDICT has been operational since Fall 2008.

The purpose of this system is to:

- Provide a central repository, accessible through a web-based portal that catalogs current computer network operational data.
- Provide secure access to multiple sources of datasets collected as a result of use and traffic on the Internet.
- Facilitate data flow among PREDICT participants for the purpose of developing new models, technologies and products that support effective threat assessment and increase cyber security capabilities.

The Homeland Security Act of 2002 [Public Law 1007-296, §302(4)] authorizes the Science and Technology Directorate to conduct “basic and applied research, development, demonstration, testing, and evaluation activities that are relevant to any or all elements of the Department, through both intramural and extramural programs.” In exercising its responsibility under the Homeland Security Act, S&T is authorized to collect information, as appropriate, to support research and development related to improving the security of the homeland.

2. The content of PREDICT is data gathered from researchers and persons associated with

PREDICT, such as data providers, data hosts, and PREDICT application review board, , and advisory board members. It also includes metadata regarding the datasets that are made available to researchers through the PREDICT program in its efforts to build products and technologies that will better protect America's computing infrastructure.

The PREDICT program uses the data collected to track usage of the datasets, to ensure compliance with operational policies and procedures, and to evaluate the effectiveness of the PREDICT program. This use of the data collected enables the PREDICT program to provide researchers with access to various types of datasets to use in their efforts to develop solutions to provide improved security to networks, applications, and data, which will benefit all computer users and help protect the homeland.

3. Using a secure web-portal, accessible through <https://www.predict.org/>, the PCC manages a centralized repository that identifies the datasets and their sources and location, and acts as the clearinghouse and operational authority for access to and release of the data. All data input to the system is either keyed in by users (Data Providers) or migrated (via upload of an XML file). The interactive nature of the PREDICT portal and the manner in which it is programmed to generate communications with researchers and/or other documentation for PREDICT, eliminates the need for additional documentation, forms, and communications to researchers using PREDICT datasets.
4. DHS S&T and the Cyber Security Division (CSD) have coordinated with other DHS S&T divisions and is aware of research activities sponsored by other agencies, and has found no duplication of efforts in the collection of the requested information, and there are no similar forms currently available that can be used for this system. DHS is not aware of any other duplications outside of this agency.
5. The PREDICT collection assists small businesses or other small entities or individual researchers because it streamlines the information collection process for persons interacting with the PREDICT program, and it helps them electronically track

interactions with PREDICT and history of usage of PREDICT datasets.

6. If the information is not collected, DHS S&T will be unable to fulfill the objectives of the PREDICT web-portal, which is to provide secure access to the catalog of the multiple sources of Internet traffic data, or facilitate data flow among PREDICT participants.
7. The special circumstances contained in item seven of the supporting statement are not applicable to this information collection.
8. By notice in the **Federal Register** on November 23, 2011 (76 FR 72426), DHS S&T notified the public that it was requesting comments on this information collection. The notice allowed for a 60-day public comment period. No comments were received. DHS S&T then by notice in the **Federal Register** on March 5, 2012 (77 FR 13135) notified the public during a 30-day public comment period. No comments were received from the public during either period.
9. DHS S&T does not provide payments or gifts to respondents in exchange for a benefit sought.
10. The PREDICT web-portal security safeguards shall meet DHS policy requirements and are documented in the PREDICT System Security Plan (SSP). All systems are subject to monitoring consistent with applicable laws, regulations, agency policies, procedures and practices. PREDICT follows the Privacy Act of 1974 (Public Law 93-589), which mandates that personal information solicited from individuals completing Federal records and forms be kept confidential. PREDICT's Privacy Threshold Analysis (PTA) was approved April 2010 by the DHS Privacy Office and was determined the system was privacy sensitive. In accordance with the privacy ruling, the PIA titled Privacy Impact Assessment for the Protected Repository for the Defense of Infrastructure Against Cyber Threats (PREDICT) dated February 25, 2008 was drafted / approved to cover the system as well as an existing System of Records Notice (SORN) – [DHS/ALL-002 - Department of Homeland Security \(DHS\) Mailing and Other Lists System](#) November 25, 2008, 73 FR 71659. . In addition, PREDICT will be operated in accordance with the E-Government Act (P.L. 107-347), December 2002 and the Federal Information Security Management Act (P.L. 107-347, Title III), December 2002. There is no pledge of confidentiality provided to the respondents.

11. There are no questions of a sensitive nature in this information collection.

12. Estimated Annualized Burden Hours and Costs (specific breakdowns included as an attachment in ROCIS submission).

Form Name / Number 100XX	Account Request Form10029	Request a Dataset Form 10032	My Datasets Page Form 10033
No. of Respondents	45	15	30
No. of Responses per Respondent	1	1	1
Avg. Burden per Response (in hours)	.25 (15 minutes)	.25 (15 minutes)	.75 (45 minutes)
Total Annual Burden (in hours)	11.25	3.75	22.5
Average Hourly Wage Rate	\$100	\$100	\$100
Total Annual Respondent Cost	\$1125	\$375	\$2250

Form Name / Number 100XX	Memorandum of Agreement PREDICT (PCC) Coordinating Center and Researcher/User Form 10035	Memorandum of Agreement PCC and Data Provider (DP) Form 10036	Memorandum of Agreement PCC and Data Host (DH) Form 10037
No. of Respondents	15	2	1
No. of Responses per Respondent	1	1	1
Avg. Burden per Response (in hours)	1 (60 minutes)	.75 (45 minutes)	.75 (45 minutes)
Total Annual Burden (in hours)	15	1.5	.75
Average Hourly Wage Rate	\$100	\$100	\$100
Total Annual Respondent Cost	\$1500	\$150	\$75

Form Name / Number 100XX	Referring Letter Form 10040	Notice of Data Disposal Form 10042	Amendment to Research/User Agreement Form 10060
No. of Respondents	45	15	15
No. of Responses per Respondent	1	1	1
Avg. Burden per Response (in hours)	1 (60 minutes)	.25 (15 minutes)	.25 (15 minutes)
Total Annual Burden (in hours)	45	3.75	3.75
Average Hourly Wage Rate	\$100	\$100	\$100
Total Annual Respondent Cost	\$4,500	\$375	\$375

Form Name / Number 100XX	Notice of Data Access Expiration Form 10061	Dataset Submission Form 10074	Memorandum of Agreement PREDICT Coordinating Center (PCC and Researcher (Researcher Organization) Form 10075
No. of Respondents	15	15	15
No. of Responses per Respondent	1	1	15
Avg. Burden per Response (in hours)	.25 (15 minutes)	.25 (15 minutes)	1 (60 minutes)
Total Annual Burden (in hours)	3.75	3.75	15
Average Hourly Wage Rate	\$100	\$100	\$100
Total Annual Respondent Cost	\$375	\$375	\$1,500

Form Name / Number 100XX	My Datasets Request Page Form 10076	Amendment to Research Organization Memorandum of Agreement Form 10077		
No. of Respondents	15	15		
No. of Responses per Respondent	1	1		
Avg. Burden per Response (in hours)	.25 (15 minutes)	.25 (15 minutes)		
Total Annual Burden (in hours)	3.75	3.75		
Average Hourly Wage Rate	\$100	\$100		
Total Annual Respondent Cost	\$375	\$375		

**Annual Reporting Burden and Respondent Cost: The total estimated ICR Public Burden in hours is 137.25.** This figure was derived by summing the total annual burden hours from all forms. **The total annual number of respondents is 258.** This figure was derived by summing the number of respondents to each form.

**Public Cost: The total estimated annual public reporting cost is \$13,725.00.** This figure was derived by summing the estimated annual respondent costs for all forms.

13. There are no capital or start-up costs associated with this information collection. There is no fee charge of for filing and of the information collection forms. Any cost burdens to respondents as a result of this information collection are identified in Item 14.
14. **Government Cost: Estimate annual cost to the federal government in relation to this information collection is \$700,000.** This cost includes an approximate annual cost \$673,000 for staffing costs related to the collection of this information for senior level directors, computer programmers, system administrators, security and engineers. Costs also include the hosting, operations & maintenance of servers, network, and travel related

to the collected information.

15. Since the launch of the PREDICT portal in March 2008 we have had six major refinements of functionality that automated data collection processes and increased efficiency of the work flow. This update streamlines and consolidates the information collected to support the evolution of the PREDICT program, and retires several deprecated forms. Starting with DHS Form 10029, the amount of data collected has been reduced. Next DHS Form 10032 was updated to add a field for the location of the research and facilitate research team tracking. Additionally, to keep up with the changes in the PREDICT program, three new forms are needed including splitting one of the currently approved forms into two documents. Specifically, DHS Form 10035 has been divided into two forms (the second form is DHS Form 10075) to support distinguishing between organizations requesting access to PREDICT datasets, broadly, and principal researchers applying for datasets for individual research; DHS Form 10074 (Dataset Submission Form) will be used for parties external to the PREDICT program to propose datasets for availability through the PREDICT web-portal; DHS Form 10076 (My Dataset Requests) will be added to confirm receipt of datasets, request more datasets in An Approved Category; DHS Form 10077 (Amendment To Research Organization MOA) will be added to add members to the specified research team. Other changes to previously approved forms include MOA verbiage updates to DHS Forms 10036, 10037 and 10061, but no change to the data collected. Also, DHS Form 10040 will be renamed to “Referring Letter”, DHS Form 10042 will be renamed to “Notice For Certificate Of Disposal”, and DHS Form 10060 will have an additional amendment added for additional dataset tracking. Finally, DHS Forms 10038, 10039 and 10041 will no longer be used. Reporting adjustments in Items 13 and 14 reflect these changes. Specific information is attached to the ROCIS submission as a separate document.
16. DHS S&T does not intend to employ the use of statistics or the publication thereof for this information collection.

17. DHS S&T will display the expiration date of OMB approval for this information collection. The current OMB number and expiration date is displayed in the upper right corner of the web pages and PDF documents. The appropriate disclaimer and privacy notice are displayed in the footers of the Web pages and on the first pages of the PDF documents.
18. DHS S&T does not request an exception to the certification of this information collection.

**B. Collection of Information Employment Statistical Methods**

Not Applicable.