Privacy Impact Assessment
for

DHS Data Analysis Tools

DHS/ALL/PIA-055

August 5, 2016

Contact Point
Mary Peterson
Chief of Staff
Office of Intelligence and Analysis
Department of Homeland Security
202-282-9632

Reviewing Official
Jonathan R. Cantor
Acting Chief Privacy Officer
Department of Homeland Security
202-343-1717

Abstract

The Department of Homeland Security's (DHS's) Office of Intelligence & Analysis (I&A) is developing, deploying, and using Data Analysis Tools (DAT) for the Department to perform enhanced analysis of DHS data sets and other data sources available to DHS in support of its homeland security mission. This privacy impact assessment examines the privacy implications of DATs, as they will analyze data sources that contain personally identifiable information (PII). It describes the types of tools the Department may develop, how the tools will use data, what information the tools will use, how information is protected when it is used in DATs, and the oversight process for DAT deployment. Additionally, this PIA describes a prototyping environment on the classified DHS network that I&A will use to facilitate the development and testing of DATs.

Overview

Introduction

Among its many responsibilities, the Department of Homeland Security's (DHS's) Office of Intelligence & Analysis (I&A) is charged with developing and using Data Analysis Tools (DAT) to access, analyze, and disseminate information, including to provide information and support to other parts of DHS.[1] I&A continues to improve its ability to equip the Department with the intelligence and information it needs to keep the homeland safe, secure, and resilient.

In order to fulfill its homeland security missions, DHS relies on the analysis and reporting of its operational and intelligence analysts. To enhance DHS's analytical capabilities, I&A will develop a series of DATs to perform enhanced analysis of DHS data sets and other data sources that are available to DHS to support its homeland security mission. DATs are Department-wide capabilities that may be used by DHS personnel. These DATs will improve the analysis and reporting of DHS data by personnel across the Department, which in turn will help the Department fulfill its mission.

DHS is issuing this privacy impact assessment (PIA) to provide greater transparency into its intelligence and data analysis activities. These DATs permit the use of DHS data only in support of previously approved purposes, and the DATs do not provide DHS personnel with new access to DHS data. Instead, the DATs are technical tools that help DHS personnel use more effectively the data to which they already have access. DHS personnel use DATs in support of missions consistent with the purpose for which the data was originally obtained by DHS and therefore only use DATs for purposes that have already been approved and documented in applicable privacy compliance documents. Consequently, the users and uses of DHS data described in the existing System of Records Notice (SORN) or PIA for a particular program or data set remain unchanged. DAT oversight procedures, including privacy compliance, are discussed below.

[1] 6 U.S.C. § 121(d)(13) and (17).

Data Analysis Tools

I&A will deploy DATs that enable the Department to perform enhanced analysis—particularly using standard statistical or quantitative methods—of DHS data sets and other data sources available to DHS in support of its homeland security mission. These tools assist homeland security efforts by helping DHS personnel to use the data they already have more efficiently and effectively. They also assist DHS in performing data-driven analysis.

DATs are applications used in the unclassified or classified network environments that help analysts perform searches of data (e.g., querying one or more databases simultaneously) or better understand the results of their searches (e.g., by using data visualization capabilities or performing a quantitative or statistical analysis). DATs allow authorized DHS users to query one or more data sets simultaneously and structure the resulting information in a way that provides context to the data or allows analysts to more accurately interpret the data they are analyzing.

There are three types of DATs, which are described below.

- **Search Tools:** A search tool provides a user with the ability to perform searches on the data sets that the user is permitted to access. An example of a search DAT is one that allows an analyst to quickly compare two data sets to which he or she already had access, such as comparing lost and stolen passports against passports utilized by known or suspected terrorists. The Data Framework search capability[2] in the Cerberus system[3] is an example of a search DAT that allows users to perform a federated query across multiple DHS data sets within the Data Framework.
- **Exploratory Analysis Tools:** Exploratory analysis tools assist a user in understanding more about the data, so that the user can determine the next steps he or she should take in the analysis. For example, exploratory analysis tools may provide the user with descriptive statistics (e.g., average, maximum, minimum, count, and odds ratio) or graphs (e.g., box plot, dot plot, and histogram). The goal of these tools is not to make conclusions, but to describe the data in a meaningful way that allows the analyst to interpret the data more easily. An example of an exploratory analysis tool would be one that identifies the number of lost and stolen passports by country of issuance or the number of passports that were lost or stolen by month.
- **Advanced Analysis Tools:** Advanced analysis tools assist a user in interpreting the data to answer intelligence or operational questions. These tools could include inferential statistics (e.g., confidence intervals, hypothesis testing, classification, regression), entity resolution algorithms, or other data modeling capabilities (e.g., network analysis, trend analysis, geospatial analysis). An example of an advanced

[2] See DHS/ALL/PIA-046(b), DHS Data Framework, dated February 27, 2015. Available at: http://www.dhs.gov/publication/dhs-all-pia-046-b-dhs-data-framework. Previous versions are also available at: http://www.dhs.gov/publication/dhs-all-pia-046-dhs-data-framework and http://www.dhs.gov/publication/dhsallpia-046-dhs-data-framework.
[3] See DHS/ALL/PIA-046-3(b) Cerberus, dated February 27, 2015. Available at: http://www.dhs.gov/sites/default/files/publications/privacy-pia-046-3-6-cerberus-02272015.pdf.

analysis DAT would be one that plots lost and stolen passports by geographic location and facilitator, which helps an analyst to visualize the geographic locations where false passports are most often identified or encountered and the underlying networks purveying or using these passports. See Section 3.2 for additional information on advanced analysis tools.

Analysts may use one or more DATs when attempting to answer an intelligence question. Using the scenarios described above, an analyst could first use a search DAT to identify the lost or stolen passports used by known or suspected terrorists. Before the creation of the search DAT, the analyst would have compiled the data and performed the comparison manually. The use of the search DAT allows the analyst to perform the comparison more quickly and reduces human error in the comparison process. Next, the analyst could use an exploratory analysis DAT to determine whether lost or stolen passports were more common from particular countries or used more frequently in certain months of the year. Finally, the analyst could use an advanced analysis DAT to plot the networks of lost or stolen passports on the map with shifts over time. Seeing these locations on a map may help an analyst better understand where the activity is occurring and identify potential trends and facilitation networks. Before the creation of this advanced analysis DAT, the analyst would have performed the network analysis and mapping visualization in separate tools and, in some cases, would have to rely on I&A's Media Services team to produce the map.

With the use of the DATs, the resulting product could be an intelligence report or paper that shows that known or suspected terrorists are using stolen passports most frequently from country X to travel along route Y, with peak travel occurring in months A and B. The analyst could have prepared a similar report without the assistance of DATs, but the use of DATs speeds up the data analysis process by making it less manual and improves the analyst's ability to make data-driven conclusions rather than simply highlighting a few observations that may or may not constitute a trend. Consequently, DATs may also strengthen the objectivity and timeliness of DHS's analysis. Before and after the creation of the DATs, the analyst used the lost and stolen passport data for the same purpose—to identify known or suspected terrorists or the tools and tactics used by known or suspected terrorists. The DATs do not change that use; rather, the DATs strengthen the analyst's ability to use the data for the same purpose.

This PIA describes DATs at a high level to protect DHS operations and analytical sources and methods. DATs described in this PIA do not include the use of basic office productivity software (e.g., Microsoft Excel, Microsoft Access) by analysts in support of the homeland security mission. DATs are not part of the underlying databases or source information technology (IT) systems. DATs do not change any data in a source system or data set or permanently retain data. (See Section 5.1 for more information on retention.) Rather, DATs are external applications used to search or interpret existing data sets. They may be used in either the DHS unclassified network or one or both of the classified DHS networks, where both unclassified and classified information can be stored, searched, and analyzed.

DAT User Base

DATs are separate from but complementary with the DHS Data Framework.[4] Users may use DATs with either data included in the Data Framework or with data that resides on the DHS classified network but outside of the Data Framework.

For data residing on the classified network but outside of the Data Framework's Cerberus,[5] the standard Data Framework access controls cannot be applied. In these situations, access to data will be restricted to a controlled list of users who have permission to view that data. The need-to-know of these users will be validated through (1) verification that the user already has access to the data (e.g., through separately approved system or security access) or (2) appropriate project documentation (e.g., concept of operations, letter of intent) that is approved by the DHS Data Access Review Council (DARC).[6]

DATs built to directly access Data Framework data will incorporate standard Data Framework access controls. Consequently, DATs will only be used on data that the user is authorized to view in the Data Framework. For example, the access controls are applied through the Data Framework search capability[7] in Cerberus, which is a DAT that searches across multiple DHS data sets and filters out search results that a user is not authorized to see. If a search yields 20 results, but the user is only authorized to see the data from 18 of those results, then the Data Framework search capability will only display the 18 results that the user is authorized to view. The user can then apply another DAT to those 18 results.

DAT Prototyping Environment

When a user comes to I&A's technical staff with a mission requirement for a tool, the staff first create the DAT in a standalone development environment on either the unclassified or classified DHS network. This development environment allows developers to build and test the

[4]See DHS/ALL/PIA-046(b), DHS Data Framework, dated February 27, 2015. Available at: http://www.dhs.gov/publication/dhs-all-pia-046-b-dhs-data-framework. Previous versions are also available at: http://www.dhs.gov/publication/dhs-all-pia-046-dhs-data-framework and http://www.dhs.gov/publication/dhsallpia-046-dhs-data-framework.

[5] Cerberus is the Data Framework's data repository on the classified network. See DHS/ALL/PIA-046-3(b) Cerberus, dated February 27, 2015. Available at: http://www.dhs.gov/sites/default/files/publications/privacy-pia-046-3-6-cerberus-02272015.pdf.

[6] The DARC is the coordinated oversight and compliance mechanism for the review of departmental initiatives and activities involving the internal or external transfer of PII through bulk data transfers in support of the Department's national and homeland security missions. The DARC ensures such initiatives or activities comply with applicable law and adequately protect the privacy, civil rights, and civil liberties of the individuals whose information is shared through such initiatives or activities.

[7] See DHS/ALL/PIA-046(b), DHS Data Framework, dated February 27, 2015. Available at: http://www.dhs.gov/publication/dhs-all-pia-046-b-dhs-data-framework. Previous versions are also available at: http://www.dhs.gov/publication/dhs-all-pia-046-dhs-data-framework and http://www.dhs.gov/publication/dhsallpia-046-dhs-data-framework.

functionality of the tool in a separate space where they will not disrupt day-to-day production operations on the network. The development environment connects to a secure, controlled enclave, known as the "prototyping environment," in which the developer and requestor of the DAT can test the tool against live data[8] to streamline the development process and to test whether a DAT may help answer an intelligence or operational question.

Currently, developers create and test prototypes using synthetic data, if it is available. However, live data is often formatted differently. To make sure the tool is compatible with the live data, developers often have to re-write the code multiple times, which increases development timelines. The ability to test tools with live data before moving them to a production environment will reduce development timelines and provide capabilities to end users more quickly. The use of live data also allows the mission requestor to better understand whether a DAT can help answer an intelligence question, what gaps may still remain, and whether there is value pursuing the tool. Some tools may not be pursued after initial development and testing.

Live data in the prototyping environment is only used for development and testing purposes, and may not be used for operational purposes. If the testing and development process reveals information about an exigent threat, the mission requestor and I&A technical team will follow a streamlined approval process to move the tool to the production environment (i.e., where it may be used by the mission requestor) or use the results within the prototyping environment. The details of this streamlined approval process will be documented in an I&A Standard Operating Procedure, developed in coordination with the DHS Office of the General Counsel, Privacy Office, and Office for Civil Rights and Civil Liberties.

Live data must be deleted from the prototyping environment within 90 days of the original ingest of the data into the prototyping environment.

In order for mission users to apply the results from DATs to operations, the DAT must be transferred to the production environment on the unclassified or classified DHS network. As part of the transfer process, I&A IT Security staff will conduct a vulnerability and compliance assessment of the DAT.

Oversight of DATs

The DHS Privacy Office, Office for Civil Rights and Civil Liberties, Office of the General Counsel, and I&A's Intelligence Oversight team provide oversight for DATs through an internal process that allows DHS oversight offices to focus on DATs that may be higher risk, ensures DATs

[8] In accordance with Report 2013-01 Data Privacy and Integrity Advisory Committee (DPIAC) on Privacy Recommendations on the Use of Live Data in Research, Testing, or Training, DHS components take a risk-based approach regarding the use of live data. This analysis begins with a rebuttable presumption that the use of live data is not approved. As it relates to DHS I&A's development and use of DATs, the DHS Privacy Office conducted a rigorous privacy risk analysis, as recommended by DPIAC Report 2013-01, and determined that the use of live data is justified. As outlined in this PIA, DHS I&A specifically justified its need to use live data to develop DATs by: specifying its intended use, explaining why synthetic data would not suffice for its intended purpose, and describing the security and technical controls in place to mitigate risk.

are reviewed by appropriate oversight groups, and meets DHS's operational timelines. As part of this process, tools are routed through one of four paths, depending on the tool's function, its user base, and the data it is accessing. These paths are described below.

- **Data Framework Oversight Review:** DATs that (1) use Data Framework data or (2) provide a new capability for the Data Framework are reviewed by the DHS Privacy Office, Office for Civil Rights and Civil Liberties, Office of the General Counsel, and I&A's Intelligence Oversight team through the Data Framework governance processes..

- **DARC Review:** DATs that (1) use Intelligence Community IC data or (2) will be used by the DHS Intelligence Enterprise IE outside of the Data Framework are reviewed by the DHS Privacy Office, Office for Civil Rights and Civil Liberties, Office of the General Counsel, and I&A's Intelligence Oversight team through the DARC.

- **Joint Data Framework Oversight and DARC Review:** DATs that implicate both Data Framework and DARC equities are reviewed by both oversight groups, since each group has specialized expertise.

- **Intelligence Oversight and Office of General Counsel Review:** Tools that only automate existing activities that are performed manually by analysts are reviewed by the DHS Office of the General Counsel and I&A's Intelligence Oversight team. The DHS Privacy Office and the Office for Civil Rights and Civil Liberties may 'spot check' these tools to ensure that automation tools do not raise any privacy, civil rights, or civil liberties concerns. For example, a tool that extracts passport numbers from unstructured text or normalizes phone numbers (e.g., adding or removing a country code, fixing formatting issues) automates functions performed manually by analysts and does not qualify as a Search, Exploratory Analysis, or Advanced Analysis DAT.

To facilitate this review, the DHS oversight offices receive a written summary of the DAT that includes detailed information, such as:

- A description of the tool's functionality;
- Its anticipated categorization (i.e., search, exploratory analysis, advanced analysis, or multi-capability);
- The intelligence or operational question or process challenge the tool is designed to address;
- The intended outcome (i.e., mission impact) for the mission requestor;
- The data that will be used or accessed;
- A list of the anticipated DHS organizations that would use the tool;
- Whether and what types of PII the DAT will access;
- Whether the DAT will generate or create data;
- Where or how the data accessed, generated, or created will be stored;
- Whether the DAT accesses commercial or publicly available data;

- Whether the DAT performs data mining;
- Whether the DAT relies on or attempts to identify individual characteristics that are protected (e.g., nationality, gender), accesses categories of information with additional protections (e.g., asylum records), or implicates other individual rights; and
- What measures are in place to evaluate the tool's effectiveness.

With this detailed information, the DHS oversight offices will conduct an informed review of the proposed DAT to ensure compliance with legal, privacy, and civil rights and civil liberties requirements. I&A initiates the review process by providing the written summary to the oversight offices and requesting review of the new tool.

To meet the stringent operational timelines, the oversight offices have ten business days from the date of the request to review the tool. The DHS Privacy Office, Office for Civil Rights and Civil Liberties, Office of the General Counsel, or I&A's Intelligence Oversight may identify objections to the tool, confirm they have no objections, request more information (e.g., a briefing) on a tool, or request changes to a tool. If an oversight office requests additional information, that oversight office and I&A will jointly establish a deadline for provision of that information and review by the oversight office. If the tool review is not completed within ten business days of notification (or, in the event of a request for additional information, by the deadline jointly established by I&A and the oversight office requesting the information), then the Under Secretary for Intelligence and Analysis (USIA), the Principal Deputy Under Secretary for Intelligence and Analysis (PDUSIA), or the Deputy Under Secretaries for Intelligence and Analysis (DUSIAs) may approve operational use on a rolling 30-day basis until the review is complete, oversight offices raise objections to the provisional use of the tool, or an issue is elevated to senior leadership.

Under exigent circumstances, the USIA, the PDUSIA, or a DUSIA may determine that a tool must be deployed immediately to meet a defined, exigent threat. In these circumstances, the approving official will notify in writing the DHS General Counsel, Chief Privacy Officer, and Officer for Civil Rights and Civil Liberties of the exigent threat and I&A's development and deployment of the tool. The approving official will provide this notification as soon as practicable, but no later than ten business days following deployment. When I&A deploys a DAT under exigent circumstances, the oversight offices will receive the normal written summary of the DAT at the same time as the DUSIA notifies them of its deployment, and those offices will review the tool at that time, consistent with the procedures outlined above.

As noted earlier, DATs do not permit access to previously inaccessible DHS data or permit that data to be used in ways incompatible with DHS's original purpose for collecting the data, and the users and uses of that data are described in the applicable privacy documentation. To ensure compliance with applicable privacy documentation, the DHS Privacy Office identifies the applicable SORN or PIA that apply to the users and use of data. I&A's development of DATs for DHS is an enterprise-wide capability. The DHS Privacy Office and Office of the General Counsel will consult with Component privacy and legal offices, as appropriate.

The DHS Privacy Office intends to conduct a Privacy Compliance Review (PCR) of DATs within two years of the publication of this PIA. The PCR will address the program's compliance with the privacy protections outlined in this PIA or other applicable program documentation and

assess whether additional privacy protections may be needed to ensure the use of this technology sustains and does not erode privacy.

Section 1.0 Authorities and Other Requirements

1.1 What specific legal authorities and/or agreements permit and define the collection of information by the project in question?

DATs do not collect any new information from the public. DHS's authority to collect the information is documented in the DHS source IT system SORN. Similarly, DATs do not change I&A's or the DHS Components' authorities to access DHS data sets. DHS provides its employees with access to DHS data sets if those employees have a need for the data in the performance of their official duties. The applicable SORN is identified as part of the DAT oversight process.

Users must also comply with any other applicable laws or policies governing their use of data. For example, I&A users must comply with Executive Order No. 12,333, United States Intelligence Activities.[9] Among other requirements, Executive Order 12,333 places limitations on the collection, retention, and use of United States Person information.

I&A is authorized to develop DATs on behalf of the DHS IE pursuant to 6 U.S.C. § 121(d)(13) and (17), which charges the USIA with establishing and utilizing, in conjunction with the DHS Chief Information Officer, a secure communications and information technology infrastructure, including advanced analytical tools, to access, receive, and analyze data and information that provides intelligence, information analysis, and support to other elements of the Department. These Department-wide DATs are consistent with and promote I&A's carrying out these responsibilities.

1.2 What Privacy Act System of Records Notice(s) (SORN(s)) apply to the information?

There are two types of data discussed in this PIA that are subject to a SORN. The first type of data is the raw or source data—the data accessed or used through a DAT. The second type of data is the analytical results—what the analyst has learned. For example, a list of lost and stolen passports and a list of passports utilized by known or suspected terrorists can both constitute raw data, while the analyst's report about the stolen passports used by a particular terrorist constitutes analytical results.

Raw data accessed or used through a DAT remains covered by the source IT system SORN. DATs do not alter the data in the source systems. Generally speaking, DHS will use DATs to support the Department's mission to prevent terrorism and other threats to homeland security. However, the authorized use of a particular data set is described in the SORN and PIA for that

[9] See Executive Order 12333, United States Intelligence Activities, as amended, July 30, 2008. Available at: http://www.archives.gov/federal-register/codification/executive-order/12333.html.

data or program, and DATs may only use data consistent with the SORN and PIA for a particular data set. DATs do not provide a user with any access to DHS data that is not described in the applicable SORN and PIA. Furthermore, all of the policy and legal controls that apply to a particular data set are maintained when the data is used with DATs. For example, if a DAT uses Passenger Name Record (PNR) data, then the DAT must apply the requirements of the Automated Targeting System (ATS) SORN[10] and its associated PIAs,[11] as well as the U.S.-European Union PNR Agreement.[12]

DHS personnel's work products must also be covered by an appropriate SORN if they are retrieved by personal identifier. This SORN may or may not be the same SORN that covers the data from a source IT system. For example, if an I&A and a U.S. Customs and Border Patrol (CBP) analyst were both analyzing PNR data stored in Cerberus, the ATS SORN[13] would cover the PNR data used by both analysts because ATS is the source system for PNR. While the CBP analyst's work product would be covered under the ATS SORN,[14] the I&A analyst's work product would be covered under the Enterprise Records System (ERS) SORN.[15]

1.3 Has a system security plan been completed for the information system(s) supporting the project?

There is a full system security plan for the creation of the tools. As presented in the plan, the tools will undergo a security review for compliance to relevant policies and processes.

1.4 Does a records retention schedule approved by the National Archives and Records Administration (NARA) exist?

Each of the DHS data sets used by DATs has an approved NARA record schedule or an interim retention schedule described in the applicable PIA. These retention schedules are implemented when the data is used by a DAT. Retention schedules are implemented through (1) data refreshes or (2) manual deletions.

For data sets on the classified network with refreshes from the underlying source data, the

[10] See DHS/CBP-006 Automated Targeting System SORN, May 22, 2012, 77 FR 30297. Available at: http://www.gpo.gov/fdsys/pkg/FR-2012-05-22/html/2012-12396.htm.
[11] See DHS/CBP/PIA-006(b) Automated Targeting System PIA, June 1, 2012. Available at: http://www.dhs.gov/sites/default/files/publications/privacy_pia_cbp_ats006b_0.pdf.
[12] See "Agreement Between the United States of America and the European Union on the Use and Transfer of Passenger Name Records to the United States Department of Homeland Security," dated December 14, 2011. Available at: https://www.dhs.gov/sites/default/files/publications/privacy/Reports/dhsprivacy_PNR%20Agreement_12_14_2011.pdf.
[13] See DHS/CBP-006 Automated Targeting System SORN, May 22, 2012, 77 FR 30297. Available at: http://www.gpo.gov/fdsys/pkg/FR-2012-05-22/html/2012-12396.htm.
[14] See DHS/CBP-006 Automated Targeting System SORN, May 22, 2012, 77 FR 30297. Available at: http://www.gpo.gov/fdsys/pkg/FR-2012-05-22/html/2012-12396.htm.
[15] See DHS/IA-001 Enterprise Records System (ERS), 73 FR 28128 (May 15, 2008). Available at: https://www.gpo.gov/fdsys/pkg/FR-2008-05-15/html/E8-10888.htm.

data is updated or deleted with each refresh of data, so that the data on the classified network mirrors the underlying source system and adheres to the source system retention schedule. Retention periods for data in the Data Framework are enforced through refreshes of data. If Data Framework data is not refreshed in a near real-time basis, then users must validate the information in the underlying DHS source system before taking any action (e.g., writing a report).

For data outside of the Data Framework, manual deletions may be necessary because a DAT accesses data from a one-time extract or another data source that is not refreshed in a timely matter. In these situations, manual deletions will be performed in accordance with the applicable records retention schedule. As an added protection, users of these types of data sets must also validate the information in the underlying DHS source system before taking any action.

Analytic results are retained according to the applicable SORN for analysts' work products. Please see Section 5.1 for more information on the retention of analytic results.

The NARA General Records Schedules [16] 3.1, "General Technology Management Records"; 3.2, "Information Systems Security Records"; 4.1, "Records Management Records"; and 4.3, "Input Records, Output Records, and Electronic Copies," cover other electronic records created by DATs, including searches and results, audit logs, and other compliance-related documentation.

1.5 If the information is covered by the Paperwork Reduction Act (PRA), provide the OMB Control number and the agency number for the collection. If there are multiple forms, include a list in an appendix.

DATs do not collect information from individuals. Information is analyzed only from the source data sets; therefore, the provisions of the Paperwork Reduction Act of 1980, 44 U.S.C. §§ 3501-21, are not applicable. For all information maintained in the underlying data sets that is subject to the Paperwork Reduction Act, the OMB Control Numbers and the agency numbers can be found within their respective PIAs.

[16] *See* https://www.archives.gov/records-mgmt/grs/.

Section 2.0 Characterization of the Information

The following questions are intended to define the scope of the information requested and/or collected, as well as reasons for its collection.

2.1 Identify the information the project collects, uses, disseminates, or maintains.

DATs enhance DHS's ability to analyze and understand the data the Department already collects pursuant to its homeland security missions. Consequently, DATs do not collect information directly from individuals. Instead, DATs use information obtained from various data sources available to DHS, many of which collected information directly from individuals. DATs access and analyze, but do not permanently retain, the information.

2.2 What are the sources of the information and how is the information collected for the project?

DATs may access: (1) DHS-collected data sets; (2) data sets made available to DHS by U.S. Government or foreign partners, some of which may be classified; and (3) commercial or publicly available data sets. DATs do not collect information directly from individuals. These data sets may be accessed consistent with DHS's authorities and applicable privacy documentation. Nothing in this PIA expands or alters DHS's ability to collect data or access particular data sets.

2.3 Does the project use information from commercial sources or publicly available data? If so, explain why and how this information is used.

DATs may access commercial or publicly available data in two forms. First, DATs may access commercial or publicly available data that has been incorporated into a DHS system of records. This commercial or publicly available data may contain PII, and its use is documented in the source system SORN and PIA. For example, ATS uses commercial data to research individuals and cargo requiring additional screening, and this collection and use of commercial data is documented in the ATS SORN[17] and PIA.[18] Similarly, a DHS source system may incorporate publicly available social media, and the use of the data in a DAT would be covered by the source system SORN.

Second, DATs may access commercial or publicly available data for reference purposes. This reference data does not include PII and is therefore not covered by a DHS system of records. DHS uses this data to aid in DHS's ability to analyze or interpret a DHS-collected data set or a

[17] See DHS/CBP-006 Automated Targeting System SORN, May 22, 2012, 77 FR 30297. Available at: https://www.gpo.gov/fdsys/pkg/FR-2012-05-22/html/2012-12396.htm.
[18] See DHS/CBP/PIA-006(b) Automated Targeting System PIA, June 1, 2012. Available at: http://www.dhs.gov/sites/default/files/publications/privacy_pia_cbp_ats006b_0.pdf.

data set provided to DHS by a U.S. Government or foreign partner. For example, DHS may use the Official Airline Guide (OAG), which contains information about real-time and historical flight data, to provide context to DHS's own travel data sources.

2.4 Discuss how accuracy of the data is ensured.

To ensure the accuracy and integrity of data, DHS data will be obtained from the source systems or the DHS Data Framework, which obtains its data from the source systems. DATs will not alter or transform data in the source systems. In the event that DHS data is not refreshed on a near real-time basis, users must check the underlying source systems to confirm the data is still accurate before taking any action (e.g., writing a report).

Data provided to the Department by a U.S. Government or Intelligence Community (IC) partners is considered to be authoritative. If there are any questions regarding the accuracy or timeliness of the data, then recipients will work with the originating agency to confirm the information. The accuracy of commercial or publicly available data ingested into a DHS system of records will be ensured through the mechanisms documented in the PIA for the applicable source data set. Commercial or publicly available data used for reference purposes will not contain PII and will only be ingested from sources that are considered reliable by industry standards.

2.5 Privacy Impact Analysis: Related to Characterization of the Information

Privacy Risk: In some instances, the age of the data used by the DATs may reduce the effectiveness of findings or render them obsolete. For example, if the underlying source system refreshes its data hourly or daily, but the same data residing on the classified network is only refreshed monthly, then a user may miss relevant information that may have been added to the record in that time (e.g., a benefit decision, recent travel in or out of the United States).

Mitigation: This risk is partially mitigated. The ability to react and respond to threat information is dependent on the timely receipt of the data. For example, the Department cannot take the appropriate action or put in place mitigation tactics if an illicit activity has already occurred by the time DHS receives notice. DHS therefore has an operational imperative to ensure the information used in its reports and operational activities is as accurate, timely, and relevant as possible. In support of this operational imperative, users must verify information in the source system before taking action if the data is not updated in near real-time.

DHS considers data provided by external partners provided to the Department for analytical and operational purposes to be authoritative. If there are any questions regarding the accuracy of externally-provided data, recipients will work with the originating agency to confirm the information. Finally, when developing intelligence reports, papers, or other work products, DHS intelligence analysts will follow good tradecraft practices, which include documenting the source of data and assessing its timeliness and reliability.

Privacy Risk: There is a risk that the use of commercial data may result in inaccurate information being used by DHS or stored in a DHS system as analytical results.

Mitigation: As noted before, this PIA does not describe a new collection of PII by DHS. DATs will not use public or commercial information containing PII unless that information has been incorporated into a system of records. In those instances, the source system SORN and PIA apply, and the users of DATs rely on the methods the source systems use to ensure data accuracy for public or commercial information containing PII.

Additionally, no PII is permanently retained or stored in a DAT, so public or commercial information containing PII will not be permanently retained or stored outside of the source IT system unless it is part of the analytical results that are retained by a DAT user at the conclusion of his or her use. Please see Section 5.1 for more information on the retention and storage of data. DHS has a mission imperative to ensure its operational and intelligence analysis and reporting is timely, accurate, and relevant, and intelligence analysts employ tradecraft measures (e.g., cross-checking sources, noting when information is derived from commercial sources or when accuracy is uncertain) to ensure the accuracy of their analysis. If a DAT user retains public or commercial information as part of his or her results, then the results would have undergone a human review to ensure the information is as timely, accurate, and relevant as possible.

In some instances, DATs may access commercial or public information that does not contain PII for reference purposes only. In these situations, DHS uses information from public or commercial sources that are considered reliable by industry standards. Intelligence analysts must still employ good tradecraft measures (e.g., noting the information is from a commercial source).

Section 3.0 Uses of the Information

The following questions require a clear description of the project's use of information.

3.1 Describe how and why the project uses the information.

As noted in the "Overview" section, DATs are technical tools that help DHS personnel use more effectively the data to which they already have access. These DATs do not permit uses of DHS data that are incompatible with the purpose for which DHS originally collected that data, and the DATs do not provide DHS personnel with new access to DHS data. Instead, the DHS personnel only use DATs for purposes that have already been approved. Generally speaking, users will use the data to identify individuals, associations, relationships, or patterns in support of homeland security missions, as they are implemented within the Department.[19] The users and uses of DHS data described in the applicable SORN or PIA for a particular program or data set remain unchanged.

[19] The five Homeland Security missions include: Prevent terrorism and enhance security; secure and manage our borders; Enforce and administer our immigration laws; Safeguard and secure cyberspace and; Ensure resilience to disasters. For more information, see "Mission," description on the DHS website, available at http://www.dhs.gov/mission.

For example, a CBP analyst may use PNR data from ATS in a DAT to analyze the travel patterns of known and suspected terrorists. The analyst may use a search tool to identify the travel records of known or suspected terrorists and then use an advanced analysis tool to visualize the travel patterns on a map. In this scenario, the CBP analyst is using the DATs to support existing CBP uses of data to prevent, detect, investigate, and prosecute terrorist offenses and related crimes and to identify individuals who would be subject to additional questioning upon arrival or departure from the United States. These uses of travel data are already articulated in the ATS SORN[20] and PIA,[21] as is the comparison and correlation of data from various data sets. DATs are simply technical tools that enhance DHS's ability to use its existing data collections for purposes that have already been articulated in SORNs and PIAs.

Once an analyst has finished his or her analysis, information obtained from DATs may be included in intelligence cables, such as an Intelligence Information Report (IIR), finished intelligence products, such as a paper or report, or other operational work products (e.g., a list of individuals selected for additional screening). As with the results of any operational analysis or intelligence product—whether developed using a DAT or through other means (e.g., using traditional office productivity software or non-technical means)—DHS may use this analysis to inform its operational activities. For example, network analysis of DHS travel data and classified holdings of derogatory information may reveal an individual with ties to a terrorist organization who has traveled to the United States. DHS now knows that this individual has ties to terrorism and access to the United States, and DHS can use this information to inform its operations, such as: referring the individual to secondary inspection for additional screening, writing a report to inform other Government or Intelligence Community agencies of the connection, denying an immigration benefit, or determining that the individual is inadmissible to the United States when he or she tries to enter the country.

3.2 Does the project use technology to conduct electronic searches, queries, or analyses in an electronic database to discover or locate a predictive pattern or an anomaly? If so, state how DHS plans to use such results.

Yes. DHS analysts could use data in exploratory analysis or advanced analysis tools to discover or locate a predictive pattern or anomaly.

Analysts could use exploratory analysis tools such as graphs or descriptive statistics (e.g., a large standard deviation) to identify anomalies in the data that require further analysis. An outlier (or infrequent observation) could be observed in a graph. Because outliers can have a profound influence on results, they should be removed if they are not relevant to the analysis (i.e., an outlier that results from a typographic error would not be relevant; an outlier that identifies aberrant behavior as an indicator of a terrorist may be relevant).

[20] See DHS/CBP-006 Automated Targeting System SORN, May 22, 2012, 77 FR 30297. Available at: http://www.gpo.gov/fdsys/pkg/FR-2012-05-22/html/2012-12396.htm.
[21] See DHS/CBP/PIA-006 Automated Targeting System PIA, August 3, 2007, and its updates. Available at: http://www.dhs.gov/publication/automated-targeting-system-ats-update.

Analysts can use advanced analysis tools to discover predictive patterns or anomalies. Three types of advanced analysis tools and their potential uses are described below.

- **Classification models** group data based on similar attributes and could be used to identify individuals who have similar suspicious behavior patterns to known or suspected terrorists or terrorist groups. For example, DHS may use classification models to identify individuals traveling through routes known to be exploited by illicit individuals. Individuals traveling along these routes may warrant additional scrutiny when they travel to or from the United States.
- **Network analysis** is a standard data modeling technique that can produce representation of a particular network. Modeling the network may reveal how interconnected the network is or predict how information might flow throughout the network. For example, DHS may visually map the interactions of various members of a human trafficking group. The visual representation may help DHS identify the trafficking group's leader or important lieutenants and how information is likely to flow throughout the network.
- **Regression analysis** is a standard statistical practice that can describe the relationship between one or more predictor variables and a dependent variable. For example, when DHS analyzes the age range of individuals who are known to have traveled for illicit purposes, such as seeking to join a terrorist group, a regression analysis may reveal that there is a strong correlation between individuals traveling to join a terrorist group and individuals who are in a particular age range (e.g., an age range of 20-25 might be a better predictor than other age ranges for illicit travel, whereas an illicit traveler aged 50 might be an anomaly). Although the regression analysis would not tell DHS why there is a link between a particular age range and traveling to join a terrorist group, DHS could use this information to help better inform its analysis and operations.

The information obtained using these DATs assists users in either refining their analysis or formulating new searches to obtain additional information.

To the extent that these activities qualify as data mining under the Federal Agency Data Mining Reporting Act of 2007, 42 U.S.C. § 2000ee-3, they will be included in DHS's annual data mining report to Congress.

3.3 Are there other components with assigned roles and responsibilities within the system?

Yes. I&A develops or adopts the DATs. DAT capabilities will be shared across DHS. DHS personnel will have access to DATs that aid in functions consistent with their job duties. For DATs used in the DHS Data Framework, access determinations will be made based on attribute-based access controls (ABAC), which rely on user attributes such as organization or job series. If necessary, DATs can be restricted to specific user groups that have been approved by the DHS oversight offices. Please see the "DAT User Base" sub-section of the Overview for more information on DAT users.

3.4 Privacy Impact Analysis: Related to the Uses of Information

Privacy Risk: There is a risk that DHS may use DATs on data for purposes other than those for which the data was collected.

Mitigation: DATs do not give a user the ability to access new DHS data sets. Consequently, to access a data set through a DAT, a user must be able to demonstrate that he or she (1) already has access to a particular data set (e.g., through the Data Framework or otherwise) for the purpose in question or (2) is participating in a project with appropriate project documentation (e.g., concept of operations, letter of intent) that is approved by the DHS DARC.

In the first scenario, DHS has existing processes in place that verify need-to-know and privacy compliance for particular data sets. For example, a user who has access to PNR data has followed both DHS and CBP processes for receiving access to that data set, which includes a verification of the user's need-to-know, compliance with binding international agreements, etc.

In the second scenario, a user may request access to a new data set as part of a Department-level project overseen by the DHS DARC. The DARC, which includes both DHS oversight offices and Components, will verify that the use is compatible with the purpose for which the data was collected. For example, if I&A, CBP, and Immigration and Customs Enforcement (ICE) analysts are working on a special project to detect and disrupt human trafficking, an ICE analyst may need access to a CBP data set and vice versa. Through the DARC, the DHS Privacy Office would verify compliance with applicable privacy documentation, including coordination with CBP, ICE, and I&A privacy offices, as appropriate.

Privacy Risk: There is a risk that the technology used by a DAT may pose unique privacy risks.

Mitigation: DHS recognizes that some types of technology have the potential to create privacy risks. For example, the same tool that allows DHS to map lost and stolen passports could be used with another data set to engage in prohibited or inappropriate behavior (e.g., mapping individuals based on religion) based on the bias of users or designers. Consequently, DATs are developed in coordination with DHS's Office of the General Counsel to ensure they operate consistent with applicable law and policy and by the DHS Privacy Office and Office for Civil Rights and Civil Liberties[22] to ensure that they use information in a manner that appropriately protects individuals' privacy, civil rights, and civil liberties. These DHS oversight offices may stipulate that DATs are not able to perform certain functions. Additionally, if appropriate, these oversight offices may limit the use of a particular DAT to a select group of users with special training.

Privacy Risk: There is a risk that DHS will take information out of context or rely on pattern and anomaly analysis or modeling to take actions or make decisions that would affect a particular group or categories of individuals in a disproportionately negative manner.

[22] More information about DHS Office of Civil Rights and Civil Liberties' oversight of screening, vetting, and intelligence activities is available at: http://www.dhs.gov/security-intelligence-and-information-policy-section.

Mitigation: Actions or decisions are not based on a single report or opinion. They are informed by thorough research and assessments that consider multiple sources of information. The availability of additional data sources through the Data Framework,[23] information sharing agreements, or IC repositories enables users to conduct more robust and comprehensive research. Users may also receive assistance from data scientists, which reduces the risk of users' misinterpreting the data. Further, pre-existing processes within the DHS Intelligence Enterprise, such as tradecraft best practices and the review of intelligence products through manager review chains, are in place to ensure that recommendations or assessments—to include those informed by research and analysis using DATs—are sound and defensible. Finally, the use of DATs will increase DHS's ability to produce data-driven reports, which helps increase the objectivity of DHS's analysis.

Privacy Risk: Users with access to DATs may abuse their privileges by accessing data that is outside the scope of their assignment, such as performing searches on themselves, friends, relatives, or neighbors.

Mitigation: This risk is partially mitigated. Users may use DATs with data included in the Data Framework or with data that resides on the DHS classified network but outside of the Data Framework. For DATs used in the Data Framework, the Data Framework's standard auditing controls will apply. At the time of this PIA, DHS had deployed tamper-resistant audit logs, although DHS needs to develop a capability to perform analysis of these audit logs to identify potentially prohibited behavior. For data residing on the classified network but outside of the Data Framework's Cerberus, access is restricted to a limited set of trusted users.

Section 4.0 Notice

The following questions seek information about the project's notice to the individual about the information collected, the right to consent to uses of said information, and the right to decline to provide information.

4.1 How does the project provide individuals notice prior to the collection of information? If notice is not provided, explain why not.

When DHS collects information from individuals, DHS provides notice through Privacy Act Statements[24] and its publication of SORNs and PIAs.[25] DATs do not collect information from the public, and the use of DATs does not change the circumstances of or purpose for DHS's collecting information. DATs provide new technical capabilities to support DHS's existing use of its information and do not change the purpose for which DHS uses the information. Consequently, DHS does not provide a separate notice of its use of DATs at the point in which DHS collects

[23] See Data Framework Privacy Impact Assessment; http://www.dhs.gov/sites/default/files/publications/privacy-pia-046b-dhs-data-framework-20150227.pdf; March 27, 2015.
[24] Pursuant to 5 U.S.C. § 552a(e)(3) agencies are required to provide what is commonly referred to as a Privacy Act Statement to all persons asked to provide personal information about themselves if that information will go into a system of records (i.e., the information will be stored and retrieved using the individual's name or other personal identifier such as a Social Security number).
[25] All DHS SORNs and PIAs are available on the DHS Privacy Office website at www.dhs.gov/privacy.

information from individuals.

4.2 What opportunities are available for individuals to consent to uses, decline to provide information, or opt out of the project?

DATs use data from DHS's existing data sets and do not collect information from individuals. Consequently, the source system data set defines an individual's opportunity to consent to uses or decline to provide information, and individuals do not have a separate opportunity to consent to the use of their data in DATs.

4.3 Privacy Impact Analysis: Related to Notice

Each of the source data systems has its own notice requirements and mechanisms for providing public notice. As explained above, no separate notice is provided to the public regarding the use of DATs. Because the DATs do not change DHS's collection or use of the information it receives from the public, the privacy impacts on the public stem from DHS's original collection and use of the data, and there are not any new privacy impacts related to notice that arise from the use of DATs. For example, the Privacy Act Statement for the Electronic System for Travel Authorization (ESTA)[26] informs applicants that DHS solicits information:

> "to determine the eligibility of, and whether there exists a law enforcement or security risk in permitting, the alien to travel to the United States. Upon review of such biographical information, the Secretary of Homeland Security shall determine whether the alien is eligible to travel to the United States under the program."[27]

The ESTA SORN also notes that "[t]he information provided through ESTA is also vetted—along with other information that the Secretary of Homeland Security determines is necessary, including information about other persons included on the ESTA application—against various security and law enforcement databases to identify those applicants who pose a security risk to the United States." DHS provides notice, both at the point of collection and through the ESTA SORN and PIA,[28] that the U.S. Government is vetting the individual's application to determine whether he or she poses a security risk to the United States. Whether this vetting involves the manual plotting of terrorist networks on network analysis charts or the use of an advanced analysis tool, the authorized purpose (i.e., determining eligibility to travel to the United States) remains the same.

[26] See DHS/CBP-009 Electronic System for Travel Authorization (ESTA), 81 FR 39680 (June 17, 2016). Available at: https://www.regulations.gov/document?D=DHS-2016-0029-0001.
[27] See "ESTA Privacy Act Statement." Available at: https://esta.cbp.dhs.gov/esta/application.html?execution=e1s1.
[28] See DHS/CBP/PIA-007 Electronic System for Travel Authorization (ESTA), dated June 2, 2008, and associated updates. Available at: https://www.dhs.gov/publication/electronic-system-travel-authorization.

Section 5.0 Data Retention by the project

The following questions are intended to outline how long the project retains the information after the initial collection.

5.1 Explain how long and for what reason the information is retained.

DATs do not permanently retain raw data accessed through a source system. Data used in the DATs will be internally "cached" (i.e., temporarily stored) in order for the information to be analyzed or processed by the tool. Generally speaking, all cached data will be purged when the user closes the DAT. However, it is important to note that some tools may run continuously in the background and therefore the data remains temporarily stored. For example, an analyst may receive alerts that new information is available about a person of interest. The implementation of retention schedules through (1) data refreshes or (2) manual deletions is described in Section 1.4.

A user may choose to retain the results of his or her analysis. If a DAT user retains results, he or she will do so consistent with the applicable SORN for his or her work product. For example, the ERS SORN[29] applies to any I&A user's work product or user output from the DATs. During its review of tools, the DHS Privacy Office, in coordination with Component privacy offices as appropriate, will identify the applicable SORN that covers the retention of any results and document the SORN in the written tool summary. Typically, users will maintain the output of the tools (such as electronic results or written analysis) in a shared space (e.g., access-controlled SharePoint sites) in which users may collaborate with other DHS users or other U.S. Government partners. This storage of results must also be consistent with the SORN that covers the user's analytical results.

Live data may be maintained in the prototyping environment for development and testing purposes only, except as noted in the "DAT Development Environment" sub-section of the Overview. Live data must be manually deleted from the prototyping environment within 90 days of the original ingest of the data into the prototyping environment. This temporary retention allows I&A to have the data available to address any user-requested modifications to the tool that come in shortly after deploying it for use on the production environment.

5.2 Privacy Impact Analysis: Related to Retention

Privacy Risk: The data used with DATs includes unclassified or classified data sets that have been copied outside of their source systems (e.g., as in the Data Framework). Consequently, the integrity of the data is dependent on regular refreshes of data, and there is a risk that data may be maintained on the classified network longer than the source system retention period if data accessed by DATs is not refreshed in near real-time.

[29] See DHS/IA-001 Enterprise Records System (ERS), 73 FR 28128 (May 15, 2008). Available at: https://www.gpo.gov/fdsys/pkg/FR-2008-05-15/html/E8-10888.htm.

Mitigation: This risk is not specific to the use of DATs; rather, this risk applies to the general activity of copying data outside of the source IT system or data set. The full mitigation of this risk is near real-time refreshes of data through processes or projects (e.g., Data Framework, other projects on DHS's classified network) that are not described in this PIA. However, DAT users mitigate this risk by confirming the results of their analysis with the source system and data provider before taking appropriate action if a data set is not refreshed in near real-time. Furthermore, with access to data provided from multiple systems, there is a greater likelihood inconsistent data will be recognized than if users accessed individual systems.

Privacy Risk: There is a risk that analysts will retain analytical results on local computers, shared drives, collaboration spaces, or in other locations and that they will not delete those results in accordance with the applicable retention period.

Mitigation: This risk is not specific to DAT use and applies to the use of data by DHS personnel in general. This risk is partially mitigated by the annual training conducted by the I&A Intelligence Oversight Officer regarding Executive Order No. 12,333, United States Intelligence Activities, as amended July 30, 2008, which includes safeguarding information concerning U.S. Persons and reviewing constitutionally protected activities. Intelligence Oversight training is mandatory for all DHS I&A personnel (including employees, detailees, and contractors). All DHS employees are also required to complete annual privacy training.

Additionally, I&A analytical products—including information accessed by DATs—are subject to review for Analytic Standards.[30] The Standards promote the protection of privacy and civil liberties by ensuring the objectivity, timeliness, relevance, and accuracy of PII used in analytic products. If reviewers or managers question the source, timeliness, or accuracy of the data, DAT users will be asked to validate and confirm the data in question.

As part of its PCR, the DHS Privacy Office will review the mechanisms in place to determine whether additional measures are needed to ensure data is being deleted in accordance with the applicable retention period.

[30] See Intelligence Community Directive (ICD) 203, *Analytic Standards*, accessible at: http://www.dni.gov/files/documents/ICD/ICD%20203%20Analytic%20Standards.pdf, dated 2 January 2015.

Section 6.0 Information Sharing

The following questions are intended to describe the scope of the project information sharing external to the Department. External sharing encompasses sharing with other federal, state and local government, and private sector entities.

6.1 Is information shared outside of DHS as part of the normal agency operations? If so, identify the organization(s) and how the information is accessed and how it is to be used.

Only DHS personnel will be permitted to use DATs that have DHS data containing PII, and DATs do not share DHS information outside of DHS. DHS may provide DATs that do not contain PII or do not contain DHS data to external partners.

Analytic work products (e.g., intelligence assessments) that may be informed or supported by DATs will be shared with external partners consistent with Components' pre-existing information sharing and dissemination guidelines approved by DHS oversight offices. These existing processes provide that, among other things, PII included in an analytic work product will be disseminated consistent with the user's authorities, policies, and procedures, including an applicable routine use outlined in the SORN for any system of records in which the results are maintained, as well as the laws and policies governing the dissemination of the underlying information provided by the source system.

To ensure that results are disseminated in a manner consistent with the SORN in which the results are maintained, the written summary of the DAT will identify the system of records for the data user in which the results of their use apply. For example, I&A personnel will share analytic work products consistent with the ERS SORN,[31] and CBP personnel will share analytic work products consistent with the ATS SORN.[32] DATs will include instructions for users regarding information sharing.

6.2 Describe how the external sharing noted in 6.1 is compatible with the SORN noted in 1.2.

No raw data from source systems is shared. Only work products may be shared, and the DAT must identify the SORN which applies to the user's work products. For example, I&A personnel's work products are covered by the ERS SORN,[33] which includes routine uses such as Routine Use D, which permits the disclosure of information "[t]o a Federal, State, local, tribal, or territorial government or agency lawfully engaged in the collection of intelligence (including national intelligence, foreign intelligence, and counterintelligence), counterterrorism, homeland

[31] See DHS/IA-001 Enterprise Records System (ERS), 73 FR 28128 (May 15, 2008). Available at: https://www.gpo.gov/fdsys/pkg/FR-2008-05-15/html/E8-10888.htm.
[32] See DHS/CBP-006 Automated Targeting System SORN, May 22, 2012, 77 FR 30297. Available at: http://www.gpo.gov/fdsys/pkg/FR-2012-05-22/html/2012-12396.htm.
[33] See DHS/IA-001 Enterprise Records System (ERS), 73 FR 28128 (May 15, 2008). Available at: https://www.gpo.gov/fdsys/pkg/FR-2008-05-15/html/E8-10888.htm.

security, law enforcement or law enforcement intelligence, and other information, where disclosure is undertaken for intelligence, counterterrorism, homeland security, or related law enforcement purposes, as authorized by U.S. Law or Executive Order, and in accordance with applicable disclosure policies." If I&A prepares an intelligence product on terrorist networks using lost or stolen passports to travel to the United States that is informed or supported by a DAT and shares this product with another federal intelligence agency for counterterrorism purposes, then Routine Use D would cover this sharing. This sharing is compatible with the purposes articulated in the ERS SORN[34] because it advances I&A's mission to "identify and assess the nature and scope of terrorist threats to the homeland"[35] and "detect and identify threats of terrorism against the United States,"[36] and would be support its responsibility "[t]o disseminate, as appropriate, information analyzed by the Department within the Department, to other agencies of the Federal Government with responsibilities relating to homeland security."[37]

Similarly, if an analyst at CBP's National Targeting Center uses a DAT to prepare a list of terrorists using lost or stolen passports, the analyst's work product is covered by the ATS SORN, [38] which includes routine uses such as Routine Use H, which permits disclosure of information "[t]o federal and foreign government intelligence or counterterrorism agencies or components where DHS becomes aware of an indication of a threat or potential threat to national or international security, or to assist in anti-terrorism efforts." If CBP writes an intelligence product informing another federal counterterrorism agency that a known or suspected terrorist on the list has traveled to the United States using a lost or stolen passport, then Routine Use H would cover this sharing. This sharing is compatible with the ATS SORN because the it supports CBP's efforts to "perform targeting of individuals who may pose a risk to border security or public safety, may be a terrorist or suspected terrorist, or may otherwise be engaged in activity in violation of U.S. law" and supports the "enforcement of the laws enforced or administered by DHS, including those related to counterterrorism."[39]

6.3 Does the project place limitations on re-dissemination?

Because only work products may be disseminated, only work products may be re-disseminated. Restrictions on re-dissemination of work products will be imposed consistent with the existing processes. These processes address a variety of policy and legal requirements, such as the proper handling of information protected by statue or regulation, such as information about asylum records[40] or victims of certain qualifying crimes.[41]

---

[34] See DHS/IA-001 Enterprise Records System (ERS), 73 FR 28128 (May 15, 2008). Available at: https://www.gpo.gov/fdsys/pkg/FR-2008-05-15/html/E8-10888.htm.
[35] 6 U.S.C. § 121(d)(1)(A).
[36] 6 U.S.C. § 121(d)(1)(B).
[37] 6 U.S.C. § 121(d)(8).
[38] See DHS/CBP-006 Automated Targeting System SORN, May 22, 2012, 77 FR 30297. Available at: http://www.gpo.gov/fdsys/pkg/FR-2012-05-22/html/2012-12396.htm.
[39] See DHS/CBP-006 Automated Targeting System SORN, May 22, 2012, 77 FR 30297. Available at: http://www.gpo.gov/fdsys/pkg/FR-2012-05-22/html/2012-12396.htm.
[40] 8 CFR § 208.6.
[41] 8 U.S.C. § 1367

Additionally, products may be subject to dissemination controls related to classified national security information, which may require that an individual has a certain security clearance to access a classified product. In all cases, individuals must establish a need-to-know to access classified information.

6.4 Describe how the project maintains a record of any disclosures outside of the Department.

DAT users will follow their existing Component processes. For Department-level initiatives, users will follow processes approved by the DHS oversight offices and outlined in appropriate project documentation.

6.5 Privacy Impact Analysis: Related to Information Sharing

Privacy Risk: There is a risk that DHS will share PII outside of the Department for a purpose that is not compatible with the purpose for which the PII was collected.

Mitigation: Results or outputs informed or supported by DATs will be shared with external partners consistent with Components' pre-existing information sharing and dissemination guidelines. For Department-level initiatives, DAT users will follow procedures approved by the Office of the General Counsel, the Privacy Office, and the Office for Civil Rights and Civil Liberties. These business rules provide that, among other things, PII included in an analytic work product will be disseminated consistent with the user's authorities, policies, and procedures, including an applicable routine use outlined in the SORN for any system of records in which the results are maintained, as well as the laws and policies governing the dissemination of the underlying information provided by the source system.

All DATs must identify the SORN that covers its output. Section 6.2 provides two examples of how a routine use could be used to share externally information that is supported by or derived from a DAT. Further, DHS provides mandatory privacy training to all employees and contractors who have access to or use PII, and all users are required to complete mandated information security training that addresses privacy as well as the proper and secure use of DHS applications.

Section 7.0 Redress

The following questions seek information about processes in place for individuals to seek redress which may include access to records about themselves, ensuring the accuracy of the information collected about them, and/or filing complaints.

7.1 What are the procedures that allow individuals to access their information?

The work product of DATs may contain classified and sensitive but unclassified information related to intelligence, counterterrorism, homeland security, and law enforcement programs. Such records are exempted from notification, access, and amendment to the extent permitted by subsection (j) and (k) of the Privacy Act, as described in the Code of Federal Regulations, and as delineated in the applicable SORN. However, the SORNs for the DHS data sets explain the procedures by which data subjects may request amendment of their information. If information is corrected or removed in the underlying source systems, its accuracy is reflected through refreshes of records in the repository or extract subject to the search capability and DATs. The procedures for individuals to address possibly inaccurate or erroneous information are described in underlying SORNs.

Individuals may seek access to their records from the underlying source system and from the system that covers that users' analytical output by following the directions set forth in the appropriate SORNs.

A request for access to non-exempt records in this system may be made by writing to the Freedom of Information Act (FOIA) Officer, Office of Intelligence and Analysis, Department of Homeland Security, Washington, DC 20528, in conformance with 6 C.F.R. Part 5, Subpart B, which provides the rules for requesting access to Privacy Act records maintained by DHS.

7.2 What procedures are in place to allow the subject individual to correct inaccurate or erroneous information?

As noted above, the work product of DATs may contain classified and sensitive but unclassified information related to intelligence, counterterrorism, homeland security, and law enforcement programs. Such records are exempted from notification, access, and amendment to the extent permitted by subsection (j) and (k) of the Privacy Act, as described in the Code of Federal Regulations, and as delineated in the applicable SORN. However, the SORNs for the DHS data sets explain the procedures by which data subjects may request amendment of their information. If information is corrected or removed in the underlying source systems, its accuracy is reflected through refreshes of records in the repository or extract subject to the search capability and DATs. The procedures for individuals to address possibly inaccurate or erroneous information are described in underlying SORNs.

Individuals may seek to amend their records in the underlying source system and in the system that covers that users' analytical output by following the directions set forth in the appropriate SORNs.

7.3 How does the project notify individuals about the procedures for correcting their information?

DATs do not collect any new information from the public. The authority to collect the information and procedures for correcting information are documented in the source IT system SORN. DHS provides general notice to the public on filing a Privacy Act request on its website at: http://www.dhs.gov/file-privacy-act-request.

7.4 Privacy Impact Analysis: Related to Redress

DATs do not collect or permanently retain any new information from the public. The SORNs for the underlying data explain the procedures by which data subjects may request amendment of their information. If information is corrected or removed in the underlying source systems, its accuracy is reflected in results returned by the search capability.

Privacy Risk: There is a risk that individuals may not have sufficient redress for exempted records.

Mitigation: Even if individuals are not able to amend exempted records, they may submit general complaints about treatment or requests for redress to the DHS Traveler Redress Inquiry Program (TRIP), 601 South 12th Street, TSA 901, Arlington, VA 22202-4220 or online at www.dhs.gov/trip.

Section 8.0 Auditing and Accountability

The following questions are intended to describe technical and policy based safeguards and security measures.

8.1 How does the project ensure that the information is used in accordance with stated practices in this PIA?

Specific auditing, accountability, and oversight measures include:

- A written summary of the DAT (described in the Overview);
- Regular review of DATs by the DHS Office for Civil Rights and Civil Liberties, DHS Privacy Office, DHS Office of the General Counsel, and I&A's Intelligence Oversight team; and
- Additional security measures described in Section 3.4.

During DAT use of both Data Framework and non-Data Framework data, log management and analysis tools will monitor and assess audit data population and network processing to identify issues related to erroneous data, false inclusion/exclusion of access and/or information, and to prove that the audit capability is immutable.

Legal and policy controls on the use and protection of information will be implemented through the policy process and integrated into the technology via access control rules enforced when a DAT directly accesses Data Framework data. This allows data protections to be built-in to the initial search and monitored enterprise-wide.

Furthermore, the DHS Privacy Office intends to conduct a PCR of DATs within two years of the publication of this PIA. The PCR will assess the program's compliance with the practices stated in this Assessment.

8.2 Describe what privacy training is provided to users either generally or specifically relevant to the project.

DHS provides mandatory privacy training to all employees and contractors who have access to or use PII, and all users are required to complete mandated information security training that addresses privacy as well as the proper and secure use of DHS applications. The Office for Civil Rights and Civil Liberties offers several training products through its Civil Liberties Institute.[42]

Persons employed by or detailed to the I&A are required to attend annual training conducted by the I&A Intelligence Oversight Officer regarding Executive Order No. 12,333, United States Intelligence Activities, as amended July 30, 2008, which includes safeguarding information concerning U.S. persons and reviewing constitutionally protected activities.

Additionally users are provided training as part of the process to gain access to certain data repositories (e.g., Data Framework) or tools. For example, all Data Framework users receive a demonstration and hands-on training that includes instructions on how to use the Data Framework search capability, how to report perceived inaccuracies, how to handle sensitive information (U.S. Person, PII, etc.) and other best practices. Software developers, data scientists, and support staff provide DAT training to new users and are available for ad hoc questions and requests.

8.3 What procedures are in place to determine which users may access the information and how does the project determine who has access?

To access a data set through a DAT, a user must be able to demonstrate that he or she (1) already has access to a particular data set (e.g., through the Data Framework or otherwise) for the purpose in question or (2) is participating in a project with appropriate project documentation (e.g., concept of operations, letter of intent) that is approved by the DHS DARC.

[42] *See* http://www.dhs.gov/civillibertiesinstitute.

All users accessing classified data will be operating at the Top Secret/Sensitive Compartmented Information classified level and must have security clearances and access approvals commensurate with that level. In addition, I&A analytic personnel will access DATs in accordance with I&A's Intelligence Oversight Procedures. The use of access controls, user PKI certificates, and the DHS user credentialing data store will ensure identity of DAT users as appropriate. If noncompliance is discovered through periodic audit reviews, appropriate disciplinary and corrective actions will be taken.

8.4 How does the project review and approve information sharing agreements, MOUs, new uses of the information, new access to the system by organizations within DHS and outside?

DATs will not share DHS PII externally and are therefore not subject to any information sharing agreements or MOUs.

Responsible Officials

Mary Peterson
Chief of Staff
Office of Intelligence and Analysis
Department of Homeland Security

Approval Signature

Original signed copy on file with the DHS Privacy Office.

Jonathan R. Cantor
Acting Chief Privacy Officer
Department of Homeland Security