# Privacy Impact Assessment

for the

# DHS Data Services Branch

**DHS Reference No. DHS/ALL/PIA-046**

**January 5, 2020**

# Abstract

The U.S. Department of Homeland Security (DHS) Data Services Branch (DSB) (previously known as Data Framework) is a Departmental program that offers data hosting, data analytics and advanced data services. The DSB provides capabilities to support advanced data architecture, data management and governance processes, and customized data services to DHS Headquarters (HQ) and operational Components to support priority missions and operational management. The DSB is chartered to establish, operate, and maintain an online analytical processing platform called Neptune. Neptune provides controlled access to DHS data in the unclassified environment across the Department. Neptune allows approved users to host, link, analyze, or share data securely and confidently across the Department.

DHS is updating and re-issuing[1] this Privacy Impact Assessment (PIA) as a result of the DHS Data Framework Act of 2018.[2] This PIA covers the overall approach and use of the program, and consolidates DHS/ALL/PIA-046 DHS Data Framework and *subsequent updates*; DHS/ALL/PIA-046-1 Neptune and *subsequent updates*; DHS/ALL/PIA-046-2 Common Entity Index Prototype and *subsequent updates*; and DHS/ALL/PIA-046-3 Cerberus and *subsequent updates*. As DHS further develops the DSB, this PIA and its Appendices will be updated.

# Overview

Section 101 of the Homeland Security Act of 2002, Pub. L. No. 107-296 (Nov. 25, 2002), as amended, established DHS as an executive department of the United States. The mission of DHS is to 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. To enable DHS to carry out this missions, the Homeland Security Act sought to eliminate information firewalls separating data held by government agencies by consolidating multiple agencies under DHS.

In spite of consolidating the various Components to establish the Department, the existing architecture of DHS databases had not been conducive to effective implementation of the "One DHS" policy.[3] Access was cumbersome, time-intensive, and required personnel to log on and query separate databases in order to determine what relevant information DHS systems contain.

[1] The Data Framework previously ingested data into the Neptune unclassified "data lake" that DHS used to receive, store, and tag data from unclassified DHS IT systems. Once tagged, the unclassified DHS data sets from Neptune were transferred to Cerberus, which was the classified data lake that DHS used to perform classified searches of unclassified DHS data sets. The Common Entity Index was an unclassified correlation engine that allowed DHS to connect disparate DHS data sets to view all available information about an identified individual. Both Cerberus and the Common Entity Index hardware have been repurposed for use in other DHS projects.

[2] Data Framework Act of 2018, Pub. L. No. 115-331 (Dec. 19, 2018).

[3] In February of 2007, the Secretary issued the DHS *Policy for Internal Information Exchange and Sharing*, referred to as the "One DHS" memorandum, to further mandate open information exchange within DHS.

As a result, the Secretary and Deputy Secretary of Homeland Security developed the DHS Data Services Branch through the Common Vetting Task Force with collaboration among the Office of the Chief Information Officer (OCIO); the Office of Strategy, Policy, and Plans (PLCY); the Office of Intelligence and Analysis (I&A); the "oversight offices" (i.e., the DHS Privacy Office (PRIV), the Office for Civil Rights and Civil Liberties (CRCL), and the Office of the General Counsel (OGC)); and operational components.

The work of the Data Services Branch was reauthorized by the DHS Data Framework Act of 2018,[4] which directs DHS to:

1. Develop a data framework to integrate existing DHS datasets and systems for access by authorized personnel in a manner consistent with relevant legal authorities and privacy, civil rights, and civil liberties protections;
2. Ensure that all information of a DHS office or component that falls within the scope of the information sharing environment, and any information or intelligence relevant to priority mission needs and capability requirements of the homeland security enterprise, is included; and
3. Ensure that the framework is accessible to DHS employees who have an appropriate security clearance, are assigned to perform a function that requires access, and are trained in applicable standards for safeguarding and using such information.

## Program Description

The DSB's primary objective is to provide a single platform for users to host, link, analyze, or share data securely and confidently across the Department. It enables the implementation of analytic capabilities and advanced data searches across DHS. The DSB provides significant benefits to DHS and its component mission operators allowing mission owners to define and develop nimble, flexible analytic capabilities to look across DHS data in the unclassified environment.

The DSB provides dynamic access controls to safeguard information from multiple sources while enhancing sharing, near real-time movement of data for operational decision-making and centralized access to data across DHS Components, as required. It also enables the effective, and efficient use and sharing of available DHS information across the DHS enterprise and, as appropriate, the U.S. Government, while protecting privacy. The DSB includes the unclassified platform Neptune which can be accessed by approved users in a manner consistent with relevant legal authorities and privacy, civil rights, and civil liberties policies and protections. Currently, the DSB is in the process of onboarding additional priority operational datasets, creating a metadata

[4] *See supra* note 1.

catalog to be shared across the enterprise, and developing analytical capabilities relevant to mission and business operator end-users.

### Data Services Branch Structure

The goal of the DSB is to provide capabilities to support advanced data architecture, governance processes, and customized data services on the unclassified network (Neptune) in a clear and accessible format.

When sharing data, the DSB defines four elements through which to control data:

1) **User attributes** — identity characteristics about the user requesting access such as organization, role, and status of required training. The DSB incorporates a role-based access control (RBAC) model, which dynamically evaluates a system user's attributes for determining fine-grained access control.

2) **Data tags** — label the data, the authoritative source system from which it originated, and the date it was ingested into Neptune. Data tagging preserves the lineage of the source data and the purpose and intent for its collection even when it is not resident in its source system. Tagging data by type allows Neptune to determine whether to grant or deny access to specific data elements and to provide the ability to release the data (i.e., data about persons who receive additional protections, such as certain special protected classes of aliens who are subject to the non-disclosure provisions of 8 U.S.C. § 1367).[5]

3) **Context** — combines the type of search and analysis (i.e., the type of query that could be performed) that can be conducted (function) with the authorized purpose for which data (i.e., core biographic, extended biographic, or encounter information – outlined below on page 4) can be used.

4) **Dynamic access control policies** — evaluate user attributes, data tags, and context to grant or deny access to DHS data in the repository based on legal authorities and appropriate policies of the Department and/or Components.

The four elements assist the DSB in creating a systematic repeatable process for providing controlled access to DHS data across the Department. As a result, the DSB ensures access to the most authoritative, timely, and accurate data available in DHS to support critical decision-making and mission functions while maintaining the authoritative source of data at the source system.

[5] Special protected classes include individuals whose information is protected by Title 8, United States Code (U.S.C.), Section 1367, confidentiality and prohibited source provisions (relating to applicants for and beneficiaries of Violence Against Women Act (VAWA), T visa, or U visa protections in accordance with 8 U.S.C. 1367(d) and Section 810 of the Violence Against Women Reauthorization Act of 2013, including VAWA self-petitioners and VAWA cancellation or removal applicants.

## Neptune

DHS uses the unclassified Neptune platform to receive, store, and tag the data from unclassified DHS IT systems. This data is sourced from multiple unclassified DHS component source IT systems which remain the authoritative standard for their data. The datasets approved for ingest into Neptune are listed in Appendix A to this PIA.

Once tagged, the unclassified DHS datasets from Neptune can be searched on the unclassified network. The DSB allows approved users the ability to access, search, and use the DSB's unclassified data within Neptune through analysis, reporting, and visualization tools, such as Tableau and Power BI, that interface with Neptune. The availability of the various searches will be limited based on need-to-know, appropriate clearance levels, and user/roles with the full range of search methodologies not necessarily authorized for each user type and/or role. These approvals are governed by the DSB oversight bodies, including PRIV, OGC, and CRCL.

Neptune currently ingests data as collected by unclassified DHS systems within a common schema that includes core biographic data, extended biographic data, and encounter data related to individuals. Core biographic data is basic biographic information that includes name, date of birth, gender, country of citizenship, and country of birth. Extended biographic data is additional biographic information about an individual that is not considered core biographic information and includes information such as address, phone number, email address, passport number, and visa number. Information, other than core or extended biographic data, that derives from a DHS law enforcement or immigration related event/process and which is collected in accordance with DHS authorities and regulation is called encounter data. The term 'encounter' is used to describe a face-to-face meeting, an electronic or paper-based transaction (such as an application for a DHS administered benefit), or information provided to the United States by a domestic or foreign government agency, aircraft operator, or other private entity. Detailed encounter data may contain DHS derogatory information, screening/vetting results, or information pertaining to third parties who help administer government programs, such as program points of contact. Derogatory information can include such information as application numbers or numbers identifying a police report.

Neptune also collects metadata, which is descriptive information characterizing the data such as name of the source system, source system identifier to allow the specific data element to be traced back to the source system, applicable retention rule along with any original system creation date, contact information for the component data provider, and date the information was ingested into Neptune. Expansion of common schema elements and refinement of data controls and tagging may occur in accordance with the governance onboarding process as approved by the Data Framework Working Group (DFWG).

The DFWG is an executive committee with a charter approved by the Secretary of Homeland Security; it is made up of the component mission operators, system owners, data stewards, other stakeholders as needed, and offices with oversight authority over the DSB to include PRIV, CRCL, and OGC. The charter defines the mission, authority, membership, responsibilities, and operating principles of the DFWG. The mission of the DFWG is to provide effective governance, oversight, coordination, and direction to the DSB and all related projects and initiatives. The DFWG also ensures the DSB's successful and timely delivery of data is in compliance with all policy requirements. The DFWG is actively involved in approving datasets for ingestion into Neptune as well as uses and user groups for the DSB.

The DSB, with the approval of the DFWG, can share data within Neptune to grant access to users who have been approved by the DFWG. The DSB allows sharing of information through secure, automated connections rather than delivering data through ad hoc transfers of portable media (e.g., discs or hard drives). The DSB plans to offer data cleansing and tagging services, within Neptune, to assist DHS components that do not have these capabilities developed.

### Data Updates, Corrections, Deletions, and Refresh

The DSB relies on the source IT systems for timely notification of updates and corrections to the datasets. Updates and corrections to the datasets will be incorporated into Neptune after all required updates to data format and structure, data tagging workbook (DTW), Neptune-hosted table schema, access control rules, and related artifacts have been completed and the updated dataset has been tested and verified through a manual data refresh in collaboration with the source IT system owner. Since such updates and corrections to datasets are not automated, there is an inherent delay in the manual propagation of such updates or corrections from the source IT system to Neptune. Any corrections or changes to the data will happen at the source IT system and will be incorporated into Neptune during the subsequent data refresh.

If a source IT system changes any of the rules, policies, or guidelines for a dataset, then the source IT system owner will be responsible for communicating those changes to the DSB so that the DSB access control rules can be updated accordingly. DSB will communicate those changes to the DFWG for its awareness.

Until the Department establishes, with approval of its oversight offices, a near real-time data refresh capability, DHS personnel will not use data from Neptune without verifying the data in the underlying source IT system. This extra step is a privacy protection that ensures data quality and an operational protection to ensure that DHS personnel are using accurate information in DHS operations.

Data will be maintained according to the retention, use, and handling provisions of the respective System of Records Notice (SORN) for those mission systems of records. Because DSB is relying on the source IT systems to notify the DSB of changes, deletions, or corrections to data, DSB will not delete data until it receives a deletion notification from the source IT system. (Note: "Deletions" will be applied as defined by the source IT system. This may mean that data is overwritten, masked, fully removed, marked as "inactive," or archived). As part of the metadata tagging process, DHS tags each dataset with a retention period, and therefore DHS can remind the underlying source IT system of the upcoming retention expiration date if the DSB has not already received a deletion notification. Additionally, the DSB can use this retention period tag to delete data in Neptune that is from IT systems that are unable to send automated delete request notifications due to a number of constraints (e.g., resources, legacy systems, disruptions to operational support).

**Uses and Users**

Neptune can be accessed only for approved uses. To add new uses, an existing user or potential user will present a new use request to the DSB through a new mission use-case supporting such request. If the DSB approves and can accommodate the new potential use, the potential user must present to the DFWG and justify the new use and demonstrate that the use is permitted under existing PIAs and SORNs. All uses must be approved by the DFWG and are listed in Appendix B to this PIA. Furthermore, the DSB will permit no new use for any dataset without specific approval for the new use from the original source IT system owner.

The Neptune system can only be accessed by approved users. While the DSB is potentially available to all DHS data users, only user groups approved by the DFWG may access the unclassified Neptune system. Approved user groups are listed in Appendix C to this PIA.

DHS provides mandatory privacy training to all employees and contractors who have access to or use personally identifiable information (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. In addition, PRIV offers role-based training for agency employees involved with information sharing and users are required to take training specific to datasets within the Neptune system. Neptune users are trained to verify information at the source system before completing any final analysis or using the information operationally. CRCL also offers several training products.

**Data Sources**

Data ingested into Neptune comes from multiple sources across DHS components. To provide greater transparency as to which DHS datasets are approved for ingestion into Neptune, Appendix A to this PIA lists all datasets approved for ingestion into Neptune.

The DSB initially planned to enforce existing source IT system data requirements, including retention rules, by relying on the source IT systems to notify the DSB of changes, deletions, or corrections to data. However, the DSB discovered that some source IT systems are not able to send such delete notifications due to a number of constraints to their systems (e.g., legacy systems that are unable to send automated update notifications). To overcome the lack of delete notifications, the DSB reviews all SORNs and PIAs for source systems; extracts a series of business rules, including retention rules; confirms those rules with the source system and the DFWG; and duplicates and enforces those rules in the DSB for those datasets that cannot supply update notifications. By ensuring that this analysis and oversight occurs when alternative retention management changes are needed, the Department will continue to ensure that risks are understood, documented, and mitigated.

The DSB prefers to receive delete notifications from the source system; however, if the DSB internally determines it must manage retention of a particular dataset, an internal retention management request is presented to the DFWG. The DSB, in collaboration with associated stakeholders, prepares and documents an analysis of the compliance rules associated with the internal retention management and presents these rules to the DFWG.

### Data Queries

To reduce the need for bulk data exports, the DSB will provide interfaces through which users can query and aggregate data. These queries may be against a single data source or multiple data sources that can be linked through common data fields. Users will typically execute a query by filling values in a web form that executes a pre-defined query against the approved data. When such pre-defined queries are insufficient for mission requirements, a user may request permission from the DFWG to further filter/aggregate/link the data through interfaces that allow the user to specify more advanced logic. In all cases, users' access to data is strictly limited to only data for which they have been approved, and all queries will be securely audited.

### Data Sharing

The DSB provides the technical capability to reduce or replace existing information sharing processes internal to DHS and other partners approved by the DFWG. The DSB, with the approval of the DFWG, can share data, including "bulk data sharing"[6] from the Neptune system. The DSB allows sharing of data through secure, automated connections rather than delivering data through ad hoc transfers of portable media (e.g., discs or hard drives).

DSB does not currently share bulk data externally; additionally, DSB does not intend, nor is authorized, to share bulk data externally.

[6] "Bulk data sharing" is defined as the transmission of large quantities of information, which, due to technical or operational considerations, is transferred without the use of discriminants reasonably likely to exclude information not relevant to the need giving rise to the recipient's request (e.g., specific identifiers, section terms). For example, transmitting a list of all ship arrivals in the US during the first quarter of last year in response to a request for such information would not be a bulk data transfer because the request and response are limited in scope to information reasonably likely to be of value to the recipient. Conversely, transmitting information about an arbitrary list of ships registered in Panama for identifying ships arriving in the United States during the first quarter of last year within that group would qualify as a bulk data transfer.

# 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?

The Department of Homeland Security Data Framework Act of 2018[7] authorizes the Secretary of Homeland Security "to provide access for appropriate personnel to law enforcement and other information of the Department, and for other purposes." Section 2 (a)(2) states, "In developing the framework required under paragraph (1), the Secretary of Homeland Security shall ensure, in accordance with all applicable statutory and regulatory requirements, the following information is included: (A) All information acquired, held, or obtained by an office or component of the Department of Homeland Security that falls within scope of the information sharing environment, including homeland security information, terrorism information, weapons of mass destruction, and national intelligence; (B) Any information or intelligence relevant to priority mission needs and capability requirements of the homeland security enterprise, as determined by the Secretary."

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

Data held within Neptune is covered by the source systems' SORNs. All approved datasets are listed in Appendix A, which includes the relevant SORN for each source system. For users of the DSB, the source system will be identified allowing verification in the source system. Appendix A will be updated as new datasets are added to the DSB.

Neptune will continue to identify the source component for all records, which enables validation in the source system. Data will be maintained according to the retention, use, and handling provisions of the respective SORNs for those source systems.

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

The DSB operates under an Authority to Operate (ATO) issued on July 16, 2018. Additionally, a system security plan for Neptune has been completed and ratified.

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

The DSB follows the retention schedule of the source systems. The DSB prefers to rely on update/delete notifications from the source systems; however, when a source system is unable to

[7] *See supra* note 1.

provide such notifications the DSB will duplicate and apply the same retention schedule as the source system.

### 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.

The DSB is exempt from the Paperwork Reduction Act (44 U.S.C. § 3510) because it does not pose questions to collect information from ten or more members of the public. However, the information maintained by the systems from which DSB sources information may be subject to the PRA.

## Section 2.0 Characterization of the Information

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

Data entering Neptune is gathered from source systems listed in Appendix A to this PIA. The data elements ingested into Neptune are listed in Appendix A along with references to the source systems' PIAs to allow transparency of both the source systems and how those source systems collect, use, disseminate, and maintain the data. These data elements include core biographic data, extended biographic data, and encounter data.

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

The DSB gathers data from other DHS source IT systems. The information in the DSB is required to come from other source systems rather than the individual because the purpose of the DSB is to remove the stove-piping of DHS IT source systems. The DSB provides centralized data access by aggregating data from multiple DHS systems, rather than being a direct gatherer or user of information.

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

The DSB does not use information from commercial sources or ingest data from publicly available sources.

### 2.4 Discuss how accuracy is ensured.

The DSB employs a robust data quality process that ensures incoming data is properly validated and cataloged. This data quality process quarantines any data received by the DSB that has incomplete data fields. The DSB can report any anomalies to the source IT system for potential correction in the source IT system. DSB depends on the accuracy and quality of data from each source system.

### 2.5 Privacy Impact Analysis: Related to Characterization of the Information

**Privacy Risk:** There is a risk that Neptune will include more datasets than are necessary to fulfill its authorized purpose.

**Mitigation:** This privacy risk is mitigated. Neptune is designed to reduce the number of copies of datasets across the Department, by creating a long-term Department-wide data solution. The number of system data delivery methods will decrease as the DSB can provide a more controlled release and transfer of information within the Department. Eventually, some data aggregation systems may be decommissioned as their capabilities are replicated and centralized within the DSB.

DHS, through the DFWG, will continue to evaluate new datasets to determine whether they are relevant and necessary to meet the purpose of the uses for the approved users of the DSB. Additional datasets will be published in Appendix A to this PIA after those datasets have been approved by the DFWG for ingestion. Further minimizing this risk, the DSB plans to prioritize the use of datasets approved for ingestion or re-ingestion listed in Appendix A, allowing for a multi-year expansion without the addition of further datasets. Finally, each dataset added to Neptune will require the development of a Privacy Threshold Analysis (PTA) to ensure formal consideration of the PIA and SORN of new source IT systems, as well as an examination of access control rules and DSB users.

**Privacy Risk:** There is a risk that DSB users will access more PII than is necessary to accomplish their specified purpose.

**Mitigation:** This privacy risk is mitigated. The DSB is designed first and foremost to support DHS users' mission needs while mitigating privacy risk. One of the hallmarks of the DSB is the ability to restrict access to types of data, including PII, within the DSB based on the user's specified purpose. To accomplish this, DSB has tagged elements from each dataset as belonging to one of three categories—core biographic, extended biographic, and encounter information—and users are only able to access the categories that are necessary to perform their function. This use of data tags allows DSB to minimize data access according to specified purpose, which is an improvement in the implementation of data minimization within DHS. The Department plans to expand the DSB's approved tagging scheme elements (i.e., common information fields used across the datasets) and refine data controls/tagging as the Department operationalizes the DSB in accordance with the governance onboarding process, as approved by the DFWG.

**Privacy Risk:** There is a risk that the DSB will encourage the replication of datasets across DHS resulting in data proliferation.

**Mitigation:** This privacy risk is mitigated. The DSB reduces the number of copies of datasets across DHS. The DSB provides a DHS-wide data solution for having a single point of sign-on and an environment for having fewer copies of datasets at various components.

**Privacy Risk:** There is a risk that PII transferred outside of the original IT system and into the DSB will not be accurate, relevant, timely, or complete.

**Mitigation:** This risk is partially mitigated. DHS developed a data quality plan that establishes a feedback mechanism to report data quality issues to source systems. In addition, all data is received unaltered, processed with data anomalies identified, logged, and reported to the source data owners for potential correction, and data quality metrics are produced for performance and compliance reporting. DSB depends on the accuracy and quality of data from each source system.

The privacy risk will not be fully mitigated until DHS develops a near real-time refresh capability. The DSB has identified timelines for manually refreshing each existing dataset and has begun implementation of limited manual dataset refreshes. For each new dataset, DHS will use the onboarding process, described above, to identify and implement manual data refresh timelines. To provide additional mitigation, DSB users will continue to be trained to understand the risk associated with data latency (due to limited refresh capabilities). Users will also be required to verify information in the source system before completing any final analysis or using the information operationally.

# Section 3.0 Uses of the Information

## 3.1 Describe how and why the project uses the information.

As described below, the DSB is approved for multiple uses, which are set forth in Appendix B to this PIA. Approved users, as listed in Appendix C, may only use the DSB for one of these approved uses. Additionally, the DSB can share bulk data with partners for the uses listed in Appendix B.

## 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.

No.

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

Yes. The purpose of the DSB is to remove the stove-piping data and to share that data with approved users across components as approved by the DFWG. All users are identified in Appendix C, which identifies those users' component or directorate.

### 3.4 <u>Privacy Impact Analysis</u>: Related to the Uses of Information

**<u>Privacy Risk</u>:** There is a risk that DHS will include data in the DSB for a purpose other than the purpose for which it was collected in the original system.

**<u>Mitigation</u>:** This privacy risk is partially mitigated. The SORNs for datasets within Neptune specify the purposes for collection, such as immigration, border security, and operational and situational awareness purposes. Each dataset is reviewed through the PTA process prior to onboarding to ensure this compatibility is sufficient. Any changes to the datasets, users, and uses will trigger a review to determine whether the purpose remains compatible and whether this risk is impacted by the addition of new datasets, uses, or users.

**<u>Privacy Risk</u>:** There is a risk that approved users will use data for purposes other than those authorized.

**<u>Mitigation</u>:** This privacy risk is mitigated. Only the DFWG can approve new users and uses with input and clearance from PRIV, CRCL, and OGC. Once a user or use is approved by the DFWG, technical controls ensure the user is only able to access data for the use or uses for which he or she has been approved, thus limiting that user's access within Neptune. Access to data is determined by a user's purpose and function, and Neptune's policy-based controls will ensure that a user is only able to access data that is permitted for a particular purpose and function.

## Section 4.0 Notice

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

The DSB does not collect information directly from individuals and, thus, cannot supply notice other than through this PIA. Because the DSB ingests datasets from DHS components, individuals are provided notice of collection as outlined in the PIAs and SORNs for those datasets. Those PIAs and SORNs are set forth in Appendix A.

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

The DSB ingests data from multiple DHS datasets. Individuals whose data is contained within the DSB have been presented with the opportunity to consent to, or deny consent to, the use of their information when that information was collected by those components and entered into the source systems, as appropriate.

### 4.3 Privacy Impact Analysis: Related to Notice

**Privacy Risk:** There is a risk that individuals may not be aware their PII is being compared with other DHS information through the DSB.

**Mitigation:** This risk is partially mitigated. The existing PIAs and SORNs for datasets incorporated into the DSB provide notice to the public that the information may be compared against other datasets and be subject to analysis for varying DHS missions. DHS will continue to review its PIAs and SORNs when new datasets, user, or capabilities are added to the DSB. However, because the DSB does not collect information directly from individuals, it is largely reliant on this PIA and the source systems' notice mechanisms.

## Section 5.0 Data Retention by the Project

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

The DSB planned to enforce existing source IT system data requirements, including retention, by relying on the source IT systems to notify the DSB of changes, deletions, or corrections to data. As contemplated, DHS would not delete any data until the DSB received a deletion notification from the source IT system. However, DHS discovered that some source IT systems are not always able to accommodate this request to send such delete notifications due to a number of constraints (e.g., resources, legacy systems, disruptions to operational support).

To address this issue, DHS developed and deployed a data management capability within the DSB that manages the source IT system's data retention rules to ensure and enforce compliance. Specifically, the DSB employs a process to capture and validate the source IT system retention rules and develop the software to enable the DSB to replicate the source IT system rules to be compliant with the rules that govern the source IT system. This process ensures that the DSB continues to follow and comply with the retention periods provided for each source IT system.

When the Department determines that it must manage retention of a particular dataset internally, the DSB presents an internal retention management request to the DFWG for formal approval. DSB collaborates with all associated stakeholders, and prepares and documents an analysis of the compliance impacts associated with the internal retention management. This documentation includes an assessment of potential risks and proposed mitigations, as well as consistency with DSB compliance protections for approval by the DFWG.

### 5.2 Privacy Impact Analysis: Related to Retention

**Privacy Risk:** There is a risk that data will be retained longer than allowed in the original source IT system.

**Mitigation:** This risk is partially mitigated. DHS has determined that the original retention period in the source IT system will apply to the data that is ingested into Neptune. The DSB uses one of two methods to ensure compliance with retention rules. The first and preferred method is for the DSB to receive notifications from the source IT system for deletion, archiving, masking, updating, or other data changes. The second method is for the DSB to duplicate the retention rules of the source IT system and to apply those retention rules internally, but only for those source IT systems that cannot supply notifications.

## Section 6.0 Information Sharing

### 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.

The DSB does not currently share bulk data externally to DHS. Bulk data sharing involves the transfer of all or portions of a dataset outside of DHS to another U.S. Government partner. Not all external sharing will qualify as bulk data sharing. For example, if DHS shares a narrowly tailored data sample (e.g., a list of ships arriving in US ports in the month of June 2018) with a partner to identify mean time between ship arrivals in peak summer, then the sample would not qualify as bulk data sharing. Whether the sharing is in bulk does not impact the oversight or controls DHS applies to external sharing, but the Department notes the distinction for transparency purposes.

Additionally, contrary to providing bulk data sharing, the DSB may consider support for non-DHS users of DSB data if required, and duly authorized, to do so in support of a mission use case. Non-DHS users will be approved by the DFWG and will be listed in Appendix C to this PIA. These non-DHS users would access the data in the DSB in the same manner as DHS users. They would be limited to using the information in a manner that is consistent with an approved use in Appendix B and only after the DFWG had approved that use.

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

DSB does not currently share bulk data externally; additionally, DSB does not intend, nor is authorized, to share bulk data externally to the Department. Notwithstanding, if DSB plans to share externally, this PIA will be updated.

### 6.3 Does the project place limitations on re-dissemination?

Since DSB does not currently share data externally and does not intend, nor is authorized, to, re-dissemination limitations are not required.

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

The DSB does not currently share data externally; additionally, DSB does not intend, nor is authorized, to share bulk data externally to the Department. If DSB is subsequently authorized to share data externally, all such external sharing, bulk or otherwise, through the DSB must be governed by an Information Sharing Services Agreement (ISSA), such as a Memorandum of Agreement. These documents shall be maintained by the DSB and contain an accounting of what records were disclosed to whom, including the date, nature, and purpose of the external sharing and the name and address of the recipient. Furthermore, all bulk data sharing shall be approved by PRIV and documented in privacy compliance documentation. This PIA will be updated if this occurs.

### 6.5 Privacy Impact Analysis: Related to Information Sharing

Because the DSB does not share data externally, there are no risks to information sharing.

## Section 7.0 Redress

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

The source system procedures for individuals to access their information are described in the respective SORNs for the source systems. The DSB relies on the accuracy of the underlying Component systems that supply the information. To the extent the current source systems collect information directly from the individual involved, the opportunity is provided for the individual to ensure the accuracy of the data submitted. An additional opportunity exists for individuals to request access to and/or correction of their record(s) in the underlying component systems, as permitted by law, DHS policy, and described in the applicable SORNs.

Individuals cannot access their information directly in Neptune, or seek redress directly with the DSB, because that could result in divergent data in DHS systems.

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

The source system procedures for individuals to address possibly inaccurate or erroneous data are described in the respective SORNs for the source systems. Furthermore, since DSB collects these datasets from source systems that provide such redress, DSB will receive updated records in Neptune when an individual corrects inaccurate or erroneous information through the applicable source systems' procedures.

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

The DSB provides notice of correction procedures through this PIA. Individuals may be provided with notification of the procedures for correcting their information in the source systems as set forth in the PIAs for source systems. DSB depends on the accuracy and quality of data from each source system. The relevant PIAs are outlined in Appendix A to this PIA.

### 7.4 Privacy Impact Analysis: Related to Redress

**Privacy Risk:** There is a risk that an individual will not be able to receive appropriate access, correction, and redress regarding the DSB's use of PII.

**Mitigation:** This risk is mitigated. The DSB collects datasets from source systems that provide redress; for such redress, individuals should reference the applicable PIAs and SORNs listed in Appendix A. Further, the Department offers redress opportunities, as appropriate through Privacy Act and Freedom of Information Act (FOIA) provisions. Individuals may contact:

Chief Privacy Officer/Chief Freedom of Information Act Officer
Department of Homeland Security
2707 Martin Luther King Jr. Avenue, S.E.
Washington, DC 20528-0628.

Requests for information are evaluated to ensure that any release of information is lawful and does not disclose information that would cause a clearly unwarranted invasion of personal privacy or that would disclose techniques and/or procedures for law enforcement investigations or prosecutions.

**Privacy Risk:** There is a risk that changes made to PII in the underlying DHS IT source system as a result of correction and redress will not be replicated in the DSB.

**Mitigation:** This risk is partially mitigated. Any source system that has the technical capability to send update notifications to the DSB can send such corrections through the regular transfer of data to the DSB. For systems without that technical capability to send update notifications, the DSB can accept redress corrections from those systems in any format in which those updates can be supplied.

## Section 8.0 Auditing and Accountability

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

A Neptune user may only use the system for an approved use. The DSB has the ability through dynamic access controls to limit access for a specific class of user to a limited set of data that meets the needs of the approved use for that user group. All determinations regarding access and approved uses are overseen and granted by the DFWG, which includes PRIV, OGC, and CRCL. For example, a user group that is approved for a "benefits and authorizations" use might not need access to all the datasets in Neptune. The DSB can limit that user group to accessing only datasets that are required for that group's approved use. Furthermore, all DSB consumers have "read-only" access. No user may edit, amend, or otherwise change source data in Neptune.

Additionally, the DSB provides for the auditing of end user system activity with the ability to log the searches performed by an end user and to sort those searches by type to better identify how an individual user is using the system.

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

Neptune users are required to take annual DHS privacy training. Additionally, users receive mandatory yearly training specific to the DSB and the datasets within the Neptune that they are approved to use. Users are trained to verify information at the source system before completing any final analysis or using the information operationally. To facilitate human review and verification at the source IT system before operational use, the Department included source system contact information in the data tagging.

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

An individual user must be a member of a user group or role that has been approved by the DFWG prior to accessing Neptune. Approved users are listed in Appendix C to this PIA. In addition to being a member of an approved user group or role, the individual user must be approved by his or her supervisor. The individual user then must demonstrate compliance with current privacy training, DSB training, and dataset specific training for all datasets to which the individual user is granted access. Only at that stage can the individual user receive access to the system.

### 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?

All ISAAs, Memoranda of Agreement/Understanding (MOA/Us), and requests for new uses and user accesses by organizations within DHS are reviewed by the program compliance team, the program director, and the DFWG in a hierarchical review and approval cycle. The program compliance team first reviews the request and recommends an approval or denial disposition to the program director. The program director then evaluates the program compliance team's assessment and recommendation and makes an informed recommendation to the DFWG. Uses and users must be approved through this process prior to updates to Appendices B and C being submitted to PRIV for review.

### 8.5 Privacy Impact Analysis: Related to Auditing and Accountability

**Privacy Risk:** There is a risk that the elements of data access control are insufficiently developed or incorrectly implemented and will fail to limit the use of the data to the authorized purposes.

**Mitigation:** This risk is mitigated. The DSB tests the user attributes and data tags to verify that controls perform correctly. The DSB will continue to audit users to ensure that controls are not circumvented.

## Responsible Officials

Eric Charles, Director
Data Services Branch Program Office
U.S. Department of Homeland Security
(202) 447-5079

Phil Letowt
Office of Chief Information Officer
U.S. Department of Homeland Security

## Approval Signature

Original, signed version on file with the DHS Privacy Office.

Dena Kozanas
Chief Privacy Officer
U.S. Department of Homeland Security
(202) 343-1717

# Appendix A: Data Services Branch Data Sets

*Last updated: March 6, 2024*

Appendix A includes details and information on approved datasets in the Data Services Branch. The dataset information in this Appendix includes: dataset name, description, relevant compliance documents, populations covered, data elements covered, data retention requirements, data refresh rates within the Data Services Branch. and the date approved to enter the Data Services Branch. As information is updated to these datasets or as datasets are added to the Data Services Branch, this Appendix will be updated accordingly.

The datasets described on the following pages are approved for the Data Services Branch. The Data Services Branch ingests data elements from these datasets. Any future changes to the elements in these datasets will be captured and updated in this Appendix. Other datasets are pending approval for the Data Services Branch. The Data Services Branch will ingest data elements from these datasets, pending approval from the Data Services Branch. governance structure, including the oversight offices and each of the dataset stewards. Any future changes to the elements in these datasets will be captured and updated in this Appendix.

## Ship Arrival Notification System (SANS)

| | |
|---|---|
| **Component** | United States Coast Guard |
| **Status** | Approved |

**Description**

The United States Coast Guard (USCG) stores Notice of Arrival and Departure (NOAD) information electronically in the SANS. The U.S. Coast Guard collects NOAD information in order to provide for the safety and security of U.S. ports and waterways and the overall security of the United States. This information allows the USCG to facilitate effectively and efficiently the entry and departure of vessels into and from the United States and assist the USCG with assigning priorities while conducting maritime safety and security missions in accordance with international and domestic regulations. PII concerning vessel owner, crew members and/or non-crew individuals is collected to give an accurate picture of who has overall responsibility for a given vessel and who is onboard that vessel. The information is collected for the purpose of ensuring the safety and security of U.S. ports and waterways and the overall security of the United States. It is used to conduct necessary screening and national security checks.

**Relevant Compliance Documents**

PIA

DHS/USCG/PIA-006(b) Vessel Requirements for the Notice of Arrival and Departure (NOAD) and Automatic Identification System (AIS) Rulemaking[8]

<u>SORN</u>

DHS/USCG-029 Notice of Arrival and Departure System of Records[9]

DHS/USCG-061 Maritime Awareness Global Network (MAGNET) System of Records[10]

**Individuals Covered**

Crew members who arrive and depart the United States by sea and individuals associated with a vessel and whose information is submitted as part of a notice of arrival or notice of departure, including but not limited to vessel owners, operators, charterers, reporting parties, 24-hour contacts, company security officers, and persons in addition to crew who arrive and depart the U.S. by sea.

**Data Elements Covered**

USCG collects information from vessels' owners, operators, masters, agents or person in charge of the vessel(s). Information is submitted at 96-hours prior to a vessel's arrival to the United States.

Notice of arrival information collected falls into the following broad categories: Vessel and Voyage Details (including arrival/departure), Crew Information, Non-Crew Information, and Cargo Information.

Specifically, the following information is collected:

Vessel and Voyage Information:

- Name of vessel;
- Name of registered owner;
- Country of registry;
- Call sign;

[8] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, UNITES STATES COAST GUARD, PRIVACY IMPACT ASSESSMENT FOR THE VESSEL REQUIREMENTS FOR NOTICES OF ARRIVAL AND DEPARTURE (NOAD) AND AUTOMATIC IDENTIFICATION SYSTEM (AIS), DHS/USCG/PIA-006 (2008 and subsequent updates), *available at* https://www.dhs.gov/privacy-documents-us-coast-guard.

[9] *See* DHS/USCG-029 Notice of Arrival and Departure System of Records, 82 Fed. Reg. 32715 (July 17, 2017), *available at* https://www.dhs.gov/system-records-notices-sorns.

[10] *See* DHS/USCG-061 Maritime Awareness Global Network (MAGNET), 73 Fed. Reg. 28143 (May 15, 2008), *available at* https://www.dhs.gov/system-records-notices-sorns.

- International Maritime Organization (IMO) international number or, if vessel does not have an assigned IMO international number, substitute with official number;
- Name of the operator;
- Name of charterer;
- Name of classification society;
- Maritime Mobile Service Identity (MMSI); and
- Vessel(s) gross tonnage.

Voyage Information:

- Arrival information:
    - Names of last five foreign ports or places visited;
    - Dates of arrival and departure for last five foreign ports or places visited;
    - For each port or place of the United States to be visited, a list of the names of the receiving facility, the port or place, the city, and the state;
    - For each port or port or place of the United States to be visited, the estimated date and time of arrival;
    - For each port or port or place in the United States to be visited, the estimated date and time of departure;
    - The location (port or place and country) or position (latitude and longitude or waterway and mile marker) of the vessel at the time of reporting;
    - The name and telephone number of a 24-hour point of contact;
    - Duration of the voyage;
    - Last five ports of call;
    - Dates of arrival and departure in last port or place visited; and
    - Estimated date and time of arrival to the entrance of port, if applicable.
- Departure information:
    - The name of departing port or place of the United States, the estimated date and time of departure;
    - Next port or place of call (including foreign), the estimated date and time of arrival; and

- The name and telephone number of a 24-hour point of contact.

Information for each crewmember onboard:

- Full name;
- Date of birth;
- Nationality;
- Identification information (type, number, issuing country, issue date, expiration date);
- Position or duties on the vessel;
- Where the crewmember embarked (list port or place and country); and
- Where the crewmember will disembark.

Information for each person onboard in addition to crew:

- Full name;
- Date of birth;
- Nationality;
- Identification information (type, number, issuing country, issue date, expiration date);
- U.S. address information;
- Where the person embarked (list port or place and country); and
- Where the person will disembark.

Cargo Information:

- A general description of cargo, other than CDC (certain dangerous cargo), onboard the vessel (e.g., grain, container, oil);
- Name of each certain dangerous cargo carried, including United Nations (UN) number, if applicable;
- Amount of each certain dangerous cargo carried;
- Operational condition of equipment required by 33 CFR 164.35;
- The date of issuance for the company's Document of Compliance certificate;
- The date of issuance of the vessel's Safety Management Certificate;

- The name of the Flag Administration, or recognized organization(s) representing the vessel flag administration, that issued those certificates International Ship and Port Facility Security Code (ISPS) Notice;
- The date of issuance for the vessels international Ship Security Certificate (ISSC), if any;
- Whether the ISSC, if any, is an initial interim ISSC, subsequent and consecutive interim ISSC, or final ISSC;
- Declaration that the approved ship security plan, if any, is being implemented;
- If a subsequent and consecutive interim ISSC, the reasons therefore;
- The name and 24-hour contact information for the Company’s Security Officer; and
- The name of the Flag Administration, or recognized security organization(s) representing the vessel flag administration, that issued the ISSC.

### Data Retention Requirements

In accordance with National Archives and Records Administration (NARA) Disposition Authority number N1-026-05-011, NOAD information on vessels and individuals maintained in the SANS is destroyed or deleted when no longer needed for reference, or after ten years, whichever is later. Outputs, which include ad-hoc reports generated for local and immediate use to provide a variety of interested parties, for example, Captain of the Port and marine safety offices, sea marshals, U.S. Customs and Border Patrol, U.S. Immigration and Customs Enforcement with the necessary information to set up security zones, scheduling boarding and inspections activities, actions for non-compliance with regulations, and other activities in support of USCG's mission to provide for safety and security of U.S. ports, are deleted after five years if they do not constitute a permanent record according to NARA.

### Data Refresh Rates within Data Services Branch

SANS data is refreshed on a near real time basis within Data Services Branch.

## Secure Flight Confirmed Matches Data

**Components** Transportation Security Administration (TSA)

**Status** Approved[11]

**Description**

The Secure Flight program matches identifying information of aviation passengers and certain non-travelers against the consolidated and integrated terrorist watch list maintained by the Federal Government in a consistent and accurate manner, while minimizing false matches and protecting personally identifiable information. Under the Secure Flight program TSA collects limited Secure Flight Passenger Data (SFPD) from certain U.S. aircraft operators and foreign carriers for the purpose of passenger watch list matching against the No Fly and Selectee list components of the Terrorist Screening Database (TSDB) or the full TSDB or other government databases, where warranted by security considerations, such as intelligence or law enforcement databases. This PIA only addresses records for individuals encountered by Secure Flight who are confirmed as matches (SF Confirmed Matches) to the TSDB. SF Confirmed Matches data does not include SFPD for passengers selected for screening based on intelligence rules.

**Relevant Compliance Documents**

PIA

DHS/TSA/PIA-018 Secure Flight Program[12]

SORN

DHS/TSA-019 Secure Flight Records[13]

**Individuals Covered**

Individuals identified in intelligence, counterintelligence, transportation security, or information system security reports and supporting materials, including but not limited to individuals involved in matters of intelligence, law enforcement or transportation security, information systems security, the compromise of classified information, or terrorism.

[11] TSA's Secure Flight Confirmed Matches dataset was approved for ingestion into the previous Data Framework and is a carryover or remnant from that prior approval that's now being ingested into Neptune.

[12] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, TRANSPORTATION SECURITY ADMINISTRATION, PRIVACY IMPACT ASSESSMENT FOR THE TSA SECURE FLIGHT PROGRAM, DHS/TSA/PIA-018 (2007 AND SUBSEQUENT UPDATES), *available at* https://www.dhs.gov/privacy-documents-transportation-security-administration-tsa.

[13] *See* DHS/TSA-019 Secure Flight Records, 80 Fed. Reg. 233 (January 5, 2015), *available at* https://www.dhs.gov/system-records-notices-sorns.

### Data Elements Covered

- Full name;
- Date of birth;
- Gender;
- Redress number (if available);
- Known traveler number (if available);
- Encounter; and
- Passport information (if available).

To manage the processing of the SFPD, TSA will require aircraft operators to include in the SFPD the following information: Reservation Control Number; Record Sequence Number; Record Type; Passenger Update Indicator; Traveler Reference Number; and Itinerary information.

### Data Retention Requirements

Pursuant to the approved National Archives and Records Administration records retention schedule N1-560-08-003-A6, routine and insignificant case files are destroyed after thirty years; significant case files are retained permanently; watch logs are destroyed after thirty years; watchlists are destroyed 99 years after date of entry or seven years after confirmation of death, whichever is sooner.

### Data Refresh Rates within Data Services Branch

Secure Flight Confirmed Matches data is refreshed on a near real-time basis within Data Services Branch.

## Statistical Data Production and Reporting Data

**Components** Office of Immigration Statistics (OIS)

**Status** Approved

### Description

The purpose of this system is to support Office of Immigration Statistics in fulfilling its mandate to regularly prepare an extensive series of analytical and statistical reports on border security, immigration enforcement activities, refugee and asylum claims, and other immigration requests and events.

### Relevant Compliance Documents

PIA

DHS/ALL/PIA-071 Office of Immigration Statistics (OIS) Statistical Data Production and Reporting[14]

SORN

DHS/ALL-045 Statistical Immigration Data Production and Reporting System of Records[15]

### Individuals Covered

Individuals and their dependents (and individuals acting on their behalf such as attorneys) interacting with the U.S. Government in its role of implementing and enforcing its immigration system and laws, including those who have applied for immigration requests or received immigration benefits, such as adjustment of status to lawful permanent resident, and those who are subject to immigration enforcement actions, including those arrested, detained, or removed from the United States for criminal or administrative violations of the Immigration and Nationality Act.

### Data Elements Covered

The specific data elements are listed below:

- Detention data, including: Location, facility, transportation information, identification numbers, book-in/book-out dates and times, custody recommendation, information about an alien's release from custody on bond,

[14] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR THE OFFICE OF IMMIGRATION STATISTICS STATISTICAL DATA PRODUCTION AND REPORTING, DHS/ALL/PIA-071 (2018), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[15] *See* DHS/ALL-045 Statistical Immigration Data Production and Reporting System of Records, 85 Fed. Reg. 14223 (March 11, 2020), *available at* https://www.dhs.gov/system-records-notices-sorns.

recognizance, or supervision, information related to prosecutorial discretion determinations, and other alerts;

- Immigration enforcement and court case-related data, including: Descriptive information of events involving alleged law violations; arrests and charges; case number; status; record number; case category; proceedings and immigration judge decisions; schedule info; court appointments; bonds; motions; appeals; disposition; case agent; date initiated and completed;
- Immigration status and history (e.g., citizenship/naturalization certificate number, removals, explanations);
- Criminal history;
- Claimed basis of eligibility for benefit(s) sought;
- Port(s) and clearance processing lane or location of crossing, secondary examination status, date(s) and time(s) of entry, status at entry(ies);
- Metadata, including: Original Row, a monotonically increasing unique identifier; Order, an integer indicating the sequential order in which events occurred; Source, the Component from which the data record originated; Dataset, the system or data repository from which the data record originated; Type 1 Variable, indicates the data type of the value stored in the element (field) named "Type 1"; Type 2 Variable, indicates the data type of the value stored in the element (field) named "Type 2"; Type 3 Variable, indicates the data type of the value stored in the element (field) named "Type 3"; Decision Variable, indicates the data type of the value stored in the element (field) named Decision

**Data Retention Requirements**

OIS has an established NARA-approved retention schedule, N1-563-09-003 (January 1, 2009), which classifies OIS records into several categories of records. Records containing PII that OIS uses to complete its statistical analyses and reporting fall into Section 6: "Research and background material used to produce the Yearbook of Immigration Statistics." The scheduled disposition provides for the data to be evaluated for remaining business need or destruction three (3) years following the end of the fiscal year in which the yearbook is produced. However, the schedule authorizes longer retention periods if records are needed for business use beyond this period. Due to many tables in the Yearbook of Immigration Statistics and accompanying reports containing tabulations of ten (10) years, the need in some cases for OIS to compare new records with records going back several decades, and the unknown nature of future requests and necessary future comparisons, a large portion of the data OIS maintains is kept for longer than three (3) years.

**Data Refresh Rates within Data Services Branch**

SDPR data is refreshed on a quarterly basis within Data Services Branch.

## United States Border Patrol Apprehensions Data

**Components** U.S. Customs and Border Protection (CBP)

**Status** Approved

**Description**

OIS currently receives the USBP Apprehensions Data Report from CBP through a manual transmission (e.g. email), performs a manual inspection and formatting, prepares the data for statistical analysis, then stores it on the OIS shared drive or USCIS' SAS Predictive Modeling Environment (PME). OIS and CBP are seeking a more efficient way to transfer, process, and securely store USBP Apprehensions report data for use by OIS. OIS has identified the DS Neptune platform as its long-term data service provider and intends to leverage Neptune's capabilities of automated ingestion, processing, and safeguarding for all OIS data.

**Relevant Compliance Documents**

PIA

DHS/ALL/PIA-071 Office of Immigration Statistics (OIS) Statistical Data Production and Reporting[16]

DHS/CBP/PIA-034 Enterprise Management Information System-Enterprise Data Warehouse (EMIS-EDW)[17]

SORN

DHS/ALL-045 Statistical Immigration Data Production and Reporting System of Records[18]

**Individuals Covered**

The USBP Apprehensions Data Report is a list of all apprehensions for people illegally entering the country between the ports of entry for a given time period (typically Fiscal Year). Their ports of entry include the coastal, northern, and southwest borders.

**Data Elements Covered**

[16] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR THE OFFICE OF IMMIGRATION STATISTICS STATISTICAL DATA PRODUCTION AND REPORTING, DHS/ALL/PIA-071 (2018), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[17] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, IMMIGRATION AND CUSTOMS ENFORCEMENT, PRIVACY IMPACT ASSESSMENT FOR THE ENTERPRISE MANAGEMENT INFORMATION SYSTEM-ENTERPRISE DATA WAREHOUSE, DHS/CBP/PIA-034 (2016), *available at* https://www.dhs.gov/privacy-documents-us-customs-and-border-protection.

[18] *See* DHS/ALL-045 Statistical Immigration Data Production and Reporting System of Records, 85 Fed. Reg. 14223 (March 11, 2020), *available at* https://www.dhs.gov/system-records-notices-sorns.

- Alien File Number
- Apprehending Site
- Apprehension Date
- Arrest Method
- Date of Birth
- Apprehended Alien's Citizenship and Country of Birth
- Criminal Conviction
- Date of Criminal Arrest
- Port of Entry and Distance from Port of Entry
- Cardinal Direction from Point of Entry
- Family Unit Type
- Gender
- Latitude/Longitude of Apprehension
- Apprehended Alien's Marital Status
- National Crime Information Center (NCIC) Code and Description
- Smuggled and Smuggled Cost
- Suspected Gang Member
- Time in U.S.
- Unaccompanied Child

**Data Retention Requirements**

To the extent information is ingested from other systems, data is retained in EMIS-EDW in accordance with the record retention requirements of those systems. For EID, data from its ENFORCE Apprehension Booking Module (EARM) is retained for 75 years to allow for reporting and long-term trend analysis."

Neptune will automatically delete or dispose of records in accordance with the designated retention policies of the source system (Note: "Deletions" will be applied as defined by the source IT system. This may mean that data is overwritten, masked, fully removed, marked as "inactive," or archived, etc.). As part of the metadata tagging process, Neptune tags each record with a retention period, and therefore can automate the deletion/disposal of the records if Neptune has not previously received a deletion notification from the program.

### Data Refresh Rates within Data Services

The USBP Apprehensions data is refreshed on a monthly basis within Data Services.

## Operation Vaccinate Our Workforce Data

**Components** Office of the Chief Information Officer (OCIO)

**Status** Approved

**Description**

The OVOW response data will be used to create visual dashboards through business intelligence services like Tableau to report staff responses to the OVOW initiative. The responses will be used to gauge a given DHS staff's desire to receive the COVID-19 vaccination, if they have already received a dose, or if they have an appointment to receive 1 of 2 doses or a single full dose.

**Relevant Compliance Documents**

PIA

DHS/ALL/PIA-088 Preventing Infectious Disease at DHS Facilities During Declared Public Health Emergencies[19]

SORN

DHS/ALL-047 Records Related to DHS Personnel, Long-Term Trainees, Contractors, Mission Support Individuals, and Visitors During a Declared Public Health Emergency System of Records[20]

**Individuals Covered**

DHS staff that respond to the ServiceNow OVOW questionnaire.

**Data Elements Covered**

The specific data elements are listed below:

- DHS Staff First Name, Last Name, Email Address
- DHS Component
- Response to Initiative (i.e., If interested in receiving a vaccination through the program)

[19] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR THE PREVENTING INFECTIOUS DISEASE AT DHS FACILITIES DURING DECLARED PUBLIC HEALTH EMERGENCIES, DHS/ALL/PIA-088 (2021), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[20] *See* DHS/ALL-047 Records Related to DHS Personnel, Long-Term Trainees, Contractors, Mission Support Individuals, and Visitors During a Declared Public Health Emergency System of Records, 85 Fed. Reg. 80127 (December 11, 2020), *available at* https://www.dhs.gov/system-records-notices-sorns.

- First Dose Date
- Second Dose Date
- Vaccine Priority Code
- Vaccination Facility Information (i.e., Nearest Facility, Facility Location, Facility Name)

**Data Retention Requirements**

Data in Neptune will be retained for 3 years in accordance with GRS 2.7, item 080 – Non-occupation Health and Program Records and will be architecture to automatically delete the records after 3 years.

**Data Refresh Rates within Data Services**

The OVOW data is refreshed three times daily (6AM, 12PM, and 6PM) within Data Services.

## Immigration Data Integration Initiative (IDII) Data Change Request - Dissemination Service (DCR-DS) Data

**Components** Office of the Chief Information Officer (OCIO)

**Status** Approved

### Description

DCR-DS will enable DHS users and systems to get current updates and information on the DHS Immigration Domain Standards that are stored in the DHS Collibra data governance center, via Neptune. In addition, the tool will help to automate and facilitate the process of synchronizing operational data repositories with authoritative DHS reference data standard assets. The service will be available via application program interface (API) calls. The service will be provided by a Neptune service layer and/or by other Data Services mechanisms.

### Relevant Compliance Documents

PIA

- N/A

SORN

- DHS/ALL-004 General Information Technology Access Account Record System of Records (GITAARS)[21]

### Individuals Covered

Data Steward(s) of the Data Standard(s) housed within Collibra.

### Data Elements Covered

1. The Email Address, First Name, Last Name of the Data Steward(s)
2. Elements within the DETL Data Standard, housed within Collibra, listed as follows:

| **DETL Data Standard**<br>A list of Detention Locations | |
|---|---|
| **Field name** | **Description** |
| Table Name | Name of Data Standard |
| Search Value | Alpha-numeric string, typically an abbreviation of the facility |
| Value Code | Alpha-numeric string, typically an abbreviation of the facility |
| Asset Id | Alpha-numeric string that uniquely-identifies an Asset |

[21] *See* DHS/ALL-004 General Information Technology Access Account Records System (GITAARS), 77 Fed. Reg. 70792 (November 27, 2012), *available at* https://www.dhs.gov/system-records-notices-sorns.

| | |
|---|---|
| Community | A grouping of zero or more Communities and Domains. A community has a globally unique name |
| Domain Type | Determines what Asset types a Domain (a logical groupings of assets) can hold |
| Domain | A logical grouping of Assets (according to their function, project, or knowledge area) |
| Description | A name descriptor that represents the business nomenclature naming of the location |
| DETL Status | Flag indicating Facility is Active or Inactive |
| Region | 3-letter code denoting a USCIS Region |
| District | 3-letter code denoting a USCIS District (UNK = unknown) |
| Statistic Code | Used to identify reporting statistics on transactions performed in a physical location that could be different than the Valid Code physical location. [S, L, C, O, F] |
| OIC Title | Additional Descriptor of the Facility |
| Building | The building number and street name of the Address |
| Room | The unit type (such as APARTMENT, SUITE, FLOOR, etc.) and Unit number |
| Street | The street name of the Address |
| City | US City, name of the jurisdictional or geographic city/town/area where the Address resides |
| State | US State, a two–letter state abbreviation |
| Zip Code | US ZIP code, a 5- or 9- digit number that represents a USPS postal code |
| DTF_BOOK_IN_PERMITTED_IND | Indicates if the facility is a "book in/out" permitted facility. |
| DTF_Close_DT | If populated, denotes the date the facility is no longer in use |
| Asset Type | Formally defines the semantics of an asset in terms of attribute types and relation types that can be instantiated for it |
| Status | Indicates whether Asset has been approved by the Data Stewart, or is a Candidate for approval, or is Archived |

## Data Retention Requirements

Neptune inherits the schedule of the source systems from which it ingests data. Within Neptune, system development records are retained in accordance with DAA-GRS 2013-0005-0007, item 011:

- Destroy 5 years after system is superseded by a new iteration, or is terminated, defunded, or no longer needed for agency/IT administrative purposes.

- The source system for DETL data standard is Collibra. Collibra stores contact information for the ICE data stewards, the info is retained in accordance with GRS 3.2, item 030:
    - Destroy when business use ceases.
- Collibra maintains a copy of the locations of immigration detention facilities, the info is retained in accordance with GRS 3.2, item 050:
    - Destroy immediately after the identical records have been captured in a subsequent backup file or at any time after the transfer request has been signed by the National Archives, but longer retention is authorized if required for business use.

**Data Refresh Rates within Data Services**

DCR-DS data is refreshed once daily within Data Services.

## DHS Hummingbird on Neptune

**Components** Office of the Chief Information Officer (OCIO)

**Status** Approved

**Description**

Hummingbird is a cloud-based case management and task tracking system for processing Afghan refugees. The application is supported by ServiceNow and performs data collection, activity tracking, data sharing, and reporting functions. The Hummingbird application was originally developed and implemented by the Department of State (DoS), but was migrated to DHS.

**Relevant Compliance Documents**

PIA

- DHS/ALL/PIA-093 Hummingbird[22]

SORN

- DHS/USCIS/ICE/CBP-001 Alien File, Index, and National File Tracking System of Record;[23]
- DHS/USCIS-007 Benefits Information System;[24]
- DHS/USCIS-010 Asylum Information and Pre-Screening System of Records;[25]
- DHS/USCIS-017 Refugee Case Processing and Security Screening Information System;[26]
- DHS/USCIS-018 Immigration Biometric and Background Check System of Records;[27]

[22] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR HUMMINGBIRD, DHS/ALL/PIA-093 (2022), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[23] *See* DHS/USCIS/ICE/CBP-001 Alien File, Index, and National File Tracking System of Records, 82 Fed. Reg. 43556 (September 18, 2017), *available at* https://www.dhs.gov/system-records-notices-sorns.

[24] *See* DHS/USCIS-007 Benefits Information System, 84 Fed. Reg. 54622 (October 10, 2019), *available at* https://www.dhs.gov/system-records-notices-sorns.

[25] *See* DHS/USCIS-010 Asylum Information and Pre-Screening System of Records, 80 Fed. Reg. 74781 (November 30, 2015), *available at* https://www.dhs.gov/system-records-notices-sorns.

[26] *See* DHS/USCIS-017 Refugee Case Processing and Security Screening Information System of Records, 81 Fed. Reg. 72075 (October 19, 2016), *available at* https://www.dhs.gov/system-records-notices-sorns.

[27] *See* DHS/USCIS-018 Immigration Biometric and Background Check (IBBC) System of Records, 83 Fed. Reg. 36950 (July 31, 2018), *available at* https://www.dhs.gov/system-records-notices-sorns.

- DHS/CBP-005 Advance Passenger Information System (APIS);[28]
- DHS/CBP-007 CBP Border Crossing Information;[29]
- DHS/CBP-011 U.S. Customs and Border Protection TECS;[30]
- DHS/CBP-021 Arrival and Departure Information System (ADIS);[31] and
- Visa Records, State-39.[32]

**Individuals Covered**

Hummingbird will collect, maintain, and use information from Afghan nationals applying for a Special Immigrant Visa (SIV) and Operation Allies refugee applicants.

**Data Elements Covered**

- Name
- Person Type (indicates if the applicant is the *Principal* or a *Derivative* of the principal)
- Relationship to Principal (if not principle)
- A-Number
- Afghan National ID (i.e., Tazkera)
- Alias
- Date of Birth
- Place of Birth
- Country of Citizenship
- Language
- Gender
- Previous Address

[28] *See* DHS/CBP-005 Advance Passenger Information System (APIS), 80 Fed. Reg. 13407 (March 13, 2015), *available at* https://www.dhs.gov/system-records-notices-sorns.
[29] *See* DHS/CBP-007 Border Crossing Information (BCI), 81 Fed. Reg. 89957 (December 13, 2016), *available at* https://www.dhs.gov/system-records-notices-sorns.
[30] *See* DHS/CBP-011 U.S. Customs and Border Protection TECS, 73 Fed. Reg. 77778 (December 19, 2008), *available at* https://www.dhs.gov/system-records-notices-sorns.
[31] *See* DHS/CBP-021 Arrival and Departure Information System (ADIS), 80 Fed. Reg. 72081 (November 18, 2015), *available at* https://www.dhs.gov/system-records-notices-sorns.
[32] *See* Visa Records, State-39, 86 Fed. Reg. 61822 (November 8, 2021), available at https://www.federalregister.gov/documents/2021/11/08/2021-24303/privacy-act-of-1974-system-of-records.

- Contact Information (e.g., mail, phone)
- Results of Security Screening (if available)
- USCIS ELIS Account Number
- Port of Entry Used
- Passport Details
- Visa Status (indicates Special Immigrant Visa applicants; this field can be left blank for non- Special Immigrant Visa applicants)
- Visa Number
- Visa Class
- Badge Number (if badges are used at safe haven)
- Self-Reported Medical Information
- Pregnancy Information
- Insurance Information (e.g., Medicaid, Marketplace)
- Dates, Adjudication Status, and Other Relevant Information for the following DHS Forms:
    - USCIS Form I-485, Application to Register Permanent Residence or Adjust Status
    - USCIS Form I-693, Report of Medical Examination and Vaccination Record
    - USCIS Form I-765, Application for Employment Authorization
    - CBP From I-94, Arrival/Departure Record

**Case Records:**

- Hummingbird Case Number (Hummingbird assigns a unique case number (HB-000000) to all cases)
- IVIS Case Number (Special Immigrant Visa's Immigrant Visa Information System number, if applicable)
- USCIS Form I-360, Petition for Amerasian, Widow(er), or Special Immigrant, Receipt Number
- Phase (auto-updates based on the completion of activities in the Activity Records)
- Processing Location and Dates (i.e., safe haven)

- Processing Delay (used to indicate if a case is experiencing a delay in processing)
- English Interpretation Required (Yes/No)
- Results of Security Screening
- Vulnerability (used to indicate instances where a case has particular protection, medical, or other vulnerability concerns that may affect processing)
- MMR Vaccination Dates
- Self-Reported Medical Condition
- Resettlement Preference
- Final Destination Airport Code
- Resettlement Agency
- Affiliate (when the Resettlement Agency has assigned the case to one of its Affiliates)
- Assurance Requested (Yes/No)
- Assurance Received (Yes/No)
- Attached Minor on Case
- View Related Cases (used to link cases in system, for example between family units)
- U.S. Point of Contact Information (e.g., name, address, contact information, relationships to individual)

**Activity Records:**

- Hummingbird Case Number
- Intake Date and Location
- Processing Date and Location
- Pre-Flight Medicals Completed (Yes/No)
- Lodging Assignment
- Self-Reported Medicals
- Vaccinations
- COVID-19 Testing
- Work Authorization

- Biometric Appointment
- Out-Processing Information
- Departure and Travel Information

### Data Retention Requirements

GRS 2.7 item 080 Non-occupational health and wellness program records: Destroy 10 years after the most recent encounter, but longer retention is authorized if needed for business use. Records older than 10 years will be deleted on biannual records clean up days. The Data Services team will implement a load procedure that automatically deletes records older than 10 years old.

### Data Refresh Rates within Data Services

The Hummingbird data is refreshed once daily within Data Services.

## Migrant Protection Protocols Case Request System

**Components** Office of the Chief Information Officer (OCIO)

**Status** Approved.

**Description**

The Migrant Protection Protocols (MPP) Case Request System Program collects data regarding requests from MPP enrollees for disenrollment from the overarching MPP program. The MPP Case Request System provides an avenue for individuals to initiate a review of their enrollment in MPP if they believe they should not be included in the program. The MPP Case Request System Program will utilize Neptune's platform to track the disenrollment of enrollees from the MPP program and provide dashboard visualizations. The MPP Case Request System is supported by the ServiceNow platform and Neptune will maintain a connection with the MPP Case Request System's ServiceNow instance to source data and generate Tableau dashboard reporting for MPP Leadership.

**Relevant Compliance Documents**

PIA

DHS/ALL/PIA-094 Migrant Protection Protocols (MPP) Case Request System[33]

SORN

DHS/USCIS/ICE/CBP-001 Alien File, Index, and National File Tracking System of Records[34]

DHS/USCIS-007 Benefits Information System[35]

DHS/CBP-011 TECS[36]

DHS/CBP-023 Border Patrol Enforcement Records (BPER)[37]

[33] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR THE MIGRANT PROTECTION PROTOCOLS (MPP) CASE REQUEST SYSTEM, DHS/ALL/PIA-094 (2022), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[34] *See* DHS/USCIS/ICE/CBP-001 Alien File, Index, and National File Tracking System of Records, 82 Fed. Reg. 43556 (September 18, 2017), *available at* https://www.dhs.gov/system-records-notices-sorns.

[35] *See* DHS/USCIS-007 Benefits Information System, 84 Fed. Reg. 54622 (October 10, 2019), *available at* https://www.dhs.gov/system-records-notices-sorns.

[36] *See* DHS/CBP-011 U.S. Customs and Border Protection TECS, 73 Fed. Reg. 77778 (December 19, 2008), *available at* https://www.dhs.gov/system-records-notices-sorns.

[37] *See* DHS/CBP-023 Border Patrol Enforcement Records (BPER), 81 Fed. Reg. 72601 (October 20, 2016), *available at* https://www.dhs.gov/system-records-notices-sorns.

DHS/CBP-006 Automated Targeting System[38]

DHS/ALL-029 Civil Rights and Civil Liberties Records[39]

**Individuals Covered**

The MPP Case Request System will collect, maintain, and use information from individuals requesting disenrollment from the MPP program.

**Data Elements Covered**

Elements within he MPP Case Request System listed below:

| Field Name | Description |
|---|---|
| elevated_for_review | Indicates if request is referred for review |
| date_referred_for_review | Indicates the date the request was referred for review |
| elevation_notes | Indicates notes after request has been elevated |
| new_int_and_circum_change | Indicate if the enrollee's circumstances changed after interview |
| disability | Indicates if enrollee has a disability |
| preferred_contact_information | Indicates preferred contact information |
| enrollee_pref_language | Indicates enrollee preferred language |
| sys_updated_on | Indicates when record was updated |
| is_there_someone_else_we_contact_about_your_concern_such_as_an_attorney_represen | Indicates if the enrollee has a legal representative |
| u_disenroll | Indicates if the enrollee has been disenrolled |
| relation_to_affected | Indicates the relationships to the noncitizen enrolled in MPP |
| uscis_explanation | Indicates the explanation from USCIS for enrollee |
| number | Indicates the unique number for request |
| review_stage | Indicates the review stage the request is in |
| sys_updated_by | Indicates when the request was updated |
| outcom_of_crcl_audit | Indicates the outcome of the CRCL audit |
| u_file_attachment_1 | Indicates if there is file attachment |
| uscis_decision_comm | Indicates when USCIS Decision Communicated Directly to Requestor (Negative Determinations) |
| safety | Indicates if there is a Safety Need |
| sys_created_on | Indicates when request was created |
| crcl_audit_of_review_decision | Indicates if CRCL has audited decision of components |
| email_address_verification | Indicates email address verification |

[38] *See* DHS/CBP-006 Automated Targeting System, 77 Fed. Reg. 30297 (May 22, 2012), *available at* https://www.dhs.gov/system-records-notices-sorns.

[39] *See* DHS/ALL-029 Civil Rights and Civil Liberties Records, 75 Fed. Reg. 39266 (July 8, 2010), *available at* https://www.dhs.gov/system-records-notices-sorns.

| | |
|---|---|
| u_no_disenroll | Indicates No Disenrollment determination |
| sys_created_by | Indicates who created the Request |
| current_location | Indicates where the MPP enrollee is located (Country, City, State) |
| cbp_review_date_completed | Indicates when CBP completed their review |
| date_submitted | Indicates when request was submitted in MPP Disenrollment Request System |
| disagree_determination | Indicates if the enrollee should not have been placed in the MPP program because of a particular vulnerability |
| pregnancy | Indicates if enrollee has a condition related to pregnancy |
| uscis_decision_communicated_to_cbp_for_service_positive_determinations | Indicates USCIS Decision Communicated to CBP for Service (Positive Determinations) |
| active | Indicates if the request is still Active |
| refer_to_crcl | Indicates if the request was referred to CRCL for Investigation |
| date_referred_to_uscis | Indicates when request was referred to USCIS |
| mental_health | Indicates if enrollee has a condition related to Mental Health |
| preparer_concent_accepted | Indicates if the Preparer’s consent is Accepted |
| uscis_review_outcome | Indicates USCIS review outcome |
| country_code | Indicates country code |
| circum_change | Indicates if circumstances have changed since I was first placed in MPP |
| chronic_illness | Indicates if enrollee has a condition related to Chronic Illness |
| work_notes | Indicates if the request has notes |
| crcl_concern_notes | Indicates if CRCL has notes related to concerns about request |
| date_referred_to_cbp | Indicates the date request was referred to CBP |
| alternate_contact_e_mail | Indicates Legal Representative E-mail |
| alternate_contact_name | Indicates Legal Representative Name |
| other_relationship_explaination | Indicates the relationship explanation to enrollee |
| date_communicated_to_requestor | Indicates Date Decision Communicated to Requestor |
| assignment_group | Indicates Assignment Group |
| component_analysis | Indicates Component Analysis |
| date_of_birth | Indicates enrollee date of birth |
| component_uploads_review_outcome | Indicate Component Uploads Review Outcome |
| systematic_concern | Indicates if there is a Potential Systemic Concern |
| date_review_completed_by_uscis | Indicates the date USCIS completed their review |
| component_analisys | Indicates component analysis notes |
| alternate_contact_phone_number | Indicates Legal Representative Phone Number |
| sys_id | Indicates unique system ID in ServiceNow |

| | |
|---|---|
| g28_submitted | Indicates if G-28 has been submitted |
| other_email | Indicates Preparer Contact Email |
| other_explanation | Indicates Other Explanation |
| country_of_birth | Indicates Enrollee Country of Birth |
| removed_from_mpp | Indicates if enrollee was removed from MPP |
| first_name | Indicates Enrollee First Name |
| assigned_to | Indicates component request is assigned to |
| related_family_reviewed_crcl | Indicates Related Family Members tab has been reviewed by CRCL |
| u_poe_details | Indicates Port of Entry Details |
| country_origin | Indicates Enrollee Citizenship |
| sys_mod_count | Indicate the number of updates that have been to request |
| date_complete | Indicates Date Closed Complete |
| last_name | Indicates Enrollee Last Name |
| need_int_or_new_fear | Indicates if they I did not previously receive an interview about my fear of remaining in Mexico |
| middle_name | Indicates Enrollee Middle Name |
| sys_tags | Indicates system tags in ServiceNow |
| detailed_concern | Indicates detailed concerned submitted by enrollee |
| other_phone | Indicates Preparer Contact Phone Number |
| alt_country_code | Indicates Legal Representative Country Code |
| fig_request | Indicates Follow-up interview |
| other_country_code | Indicates Alternate Contact Country Code |
| related_family_members_tab_has_been_reviewed | Related Family Members tab has been reviewed |
| email_address | Indicates Enrollee E-mail Address |
| date_decision_was_communicated_to_requestor | Indicates CBP Decision Communicated to Requestor |
| related_family_reviewed_uscis | Indicates Related Family Members tab has been reviewed |
| best_phone_number | Indicates Enrollee Phone Number |
| other_name | Indicates Preparer Contact Name |
| created_from_related_family | Indicates if Created from previous MPP submission |
| crcl_audit_analysis | CRCL Audit Analysis |
| legal_concent_accepted | Indicates if Legal Consent is Accepted |
| a_number | Indicates Enrollee A-Number |

## Data Retention Requirements

DHS Records Management is working to determine a records retention plan for information collected as part of the MPP Case Request System. Records will be held permanently until a records retention schedule can be determined. Once the schedule has been properly developed, the MPP Case Request System program will seek approval from the National Archives and Records Administration (NARA). Upon approval, the proposed schedule will provide retention and disposition authorities for the records of mission related MPP Case Request System materials. The MPP Case Request System program and the ServiceNow platform administrators will ensure that the retention requirements are followed.

### Data Refresh Rates within Data Services

Neptune will check and refresh data every five minutes between 8AM and 9AM with Tableau dashboard being refreshed at 9AM.

## Program Product Management (PPM)

**Components** Office of the Chief Information Officer (OCIO)

**Status** Approved

**Description**

The PPM project is designed to be a "one stop shop" for OCIO Solutions Development Directorate (SDD)[40] customers and senior executives to see all work being performed for them by SDD teams. It provides them with artifacts/documents needed to track work related information, the ability to track finances and track spending against the MOAs, and their respective links to provide access to relevant documentation to gather insights and facilitate decision-making. PPM is tasked with establishing a workflow to aggregate, transform, analyze, and report the work SDD provides to its customers. Neptune will source data from Maestro[41] (e.g., organizations, programs, products, tasks) and MOA information (e.g., MOAs, funding, spend) available in the CIO Financial Integration Tool (CFIT) system.[42]

**Relevant Compliance Documents**

PIA

DHS/ALL/PIA-059 DHS Employee Collaboration Tools[43]

SORN

DHS/ALL-004 General Information Technology Access Account Records System (GITAARS)[44]

DHS/ALL-042 Personnel Networking and Collaboration System of Records[45]

[40] Solutions Development Directorate's mission is to deliver the services required by the homeland security entrprise for mission, business management, and information technology support. SDD's focus is to provide high-quality delivery of innovative solutions to its customers within a dynamic mission, business, and information technology environment, available at https://dhsconnect.dhs.gov/org/comp/mgmt/ocio/SDD.

[41] Maestro is a DHS-approved platform that enables IT and business teams to deliver mission critical work more efficiently. Maestro provides a centralized suite of productivity tools that helps teams make informed decisions that maximize efficiency and increase collaboration, available at https://maestro.dhs.gov/.

[42] CFIT stores all financial data including MOA funding, funds available, funds expensed, and funds remaining. It captures details related to the project funding and task orders.

[43] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR THE DHS EMPLOYEE COLLABORATION TOOLS DHS/ALL/PIA-059 (2017), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[44] *See* DHS/ALL-004 General Information Technology Access Account Records System (GITAARS), 77 Fed. Reg. 70792 (November 27, 2012), *available at* https://www.dhs.gov/system-records-notices-sorns.

[45] See DHS/ALL-042 Personnel Networking and Collaboration System of Records, 83 FR 8685 (February 28, 2018), *available at* https://www.dhs.gov/system-records-notices-sorns.

## Individuals Covered

PII (e.g., name, phone, email) will be utilized to show which DHS employee/contractor has oversight for different MOAs. Dashboards will display this information, but access to this information will be limited to individuals who already have access to this information.

## Data Elements Covered

The data elements below and their corresponding descriptions will be the datasets Neptune will source and store. The datasets will then be combined to create Tableau reports to give insights on Spend Plans, MOA Development, and the DHS employee/contractor that has oversight of the MOA development.

| Field Name | Description | Data Type |
|---|---|---|
| **cust_prog Table** | | |
| Key | Indicates the unique Key ID for this record | Varchar |
| Customer_Program | Indicates the Short Name for the Customer e.g. org of system abbreviation | Varchar |
| Customer_Program_Name | Indicates the Customer Program Full Name | Varchar |
| Customer_Organization | Indicates the Customer Organization | Varchar |
| Status | Whether SDD is currently providing services to this customer (active) or no-longer providing services (inactive) | Varchar |
| SDD_Internal | Indicates if the Customer is Internal to SDD | Varchar |
| SDD_Customer | Indicates the SDD Customer | Varchar |
| MOA_Title | Indicates the MOA title | Varchar |
| MOA_Link | Provides a link to the MOA | Text |
| Systems | Tenant systems, consumer/provider systems that use SDD services | Varchar |
| Customer_Lead | Individual(s) who are the lead POC for the Customer | Varchar |
| Technical_Lead | Individual(s) who are the technical leads for the Customer | Varchar |
| Other_POCs | Other points of contact for the customer | Varchar |
| Distribution_Lists | Distributions lists that are used to communicate to the customer teams | Varchar |
| Products | Products that SDD provides to this customer | Varchar |
| Created | Indicates the time when the record was created | DateTime |
| Updated | Indicates the time when the record was updated | DateTime |
| **programs Table** | | |
| Key | Indicates the unique Key ID for this record | Varchar |
| Program | Indicates the Program Acronym | Varchar |
| Name | Indicates the Program Full Name | Varchar |
| Description | Indicates a brief description of the program its mission | Varchar |
| Status | Whether the program is active in SDD or not | Varchar |

| | | |
|---|---|---|
| Status_New | Whether the program is active in SDD or not | Varchar |
| Program_Manager | Individual who manages the program | Varchar |
| Organization | Organization that manages the program | Varchar |
| Teams | Teams that do work for the program | Varchar |
| Jira_Project | Jira projects that are used to track work for the program | Varchar |
| Systems | FISMA systems that the program manages and operates (typically the Program Manager is the System Owner) | Varchar |
| Mobius_Activity | Mobius activities that describe Program capabilities | Varchar |
| Created | Indicates the time when the record was created | Varchar |
| Updated | Indicates the time when the record was updated | DateTime |
| Key | Indicates the unique Key ID for this record | DateTime |
| **Products Table** | | |
| Key | Indicates the unique Key ID for this record | Varchar |
| Product_Name | Indicates the Product Name | Varchar |
| Product_Description | Indicates the Product Description | Varchar |
| Product_Features | Indicates the Product Features | Varchar |
| Status | Indicates the Product Status | Varchar |
| Product_Category | Indicate the Product Category | Varchar |
| Product_Owner | Indicates the Product Owner | Varchar |
| Product_O_M_Service_Lead | Indicates the Product O&M Service Lead | Varchar |
| Program | Indicates the Products Program | Varchar |
| Created | Indicates the time when the record was created | DateTime |
| Updated | Indicates the time when the record was updated | DateTime |
| **SDD_Customers Table** | | |
| Key | Indicates the unique Key ID for this record | Varchar |
| SDD_Customer | Indicates the SDD customer abbreviation | Varchar |
| SDD_Customer_Name | Indicates the SDD Customer Full Name | Varchar |
| SDD_Customer_Organization | Indicates the Customer Organization Path | Varchar |
| Customer_Portal | URL for the Customer's portal; link to main Confluence page for the Customer's portal, if one exists | Varchar |
| Created | Indicates the time when the record was created | DateTime |
| Updated | Indicates the time when the record was updated | DateTime |
| **Org Table** | | |
| Key | Indicates the unique Key ID for this record | Varchar |
| Organization | Organization name with full path ex. HQ/MGMT/OCIO/SDD | Varchar |
| Organization_Name | Full name of the organization | Varchar |
| Organization_Acronym | Organization's acronym (recommended max 5 characters) | Varchar |
| Org_Type | Type of organization e.g. Agency, DHS Component | Varchar |
| Organization_Code | DHS Organizational Code | Varchar |
| Manager | Indicates the Organizations Manager | Varchar |

| | | |
|---|---|---|
| Parent | Indicates Parent Organization | Varchar |
| Org_Level | Level of the organization in its hierarchy | Varchar |
| Level_1 | Level 1 organization in its hierarchy | Varchar |
| Level_2 | Level 2 organization in its hierarchy | Varchar |
| Level_3 | Level 3 organization in its hierarchy | Varchar |
| Level_4 | Level 4 organization in its hierarchy | Varchar |
| Level_5 | Level 5 organization in its hierarchy | Varchar |
| Level_6 | Level 6 organization in its hierarchy | Varchar |
| Created | Indicates the time when the record was created | DateTime |
| Updated | Indicates the time when the record was updated | DateTime |
| **Systems Table** | | |
| Key | Indicates the unique Key ID for this record | Varchar |
| System | Indicates the System Acronym | Varchar |
| System_Name | Indicates the System Full Name | Varchar |
| FISMA_System_Name | Indicates the Systems FISMA Name | Varchar |
| FISMA_Composite_System_ID | FISMA registered asset ID associated with the system. | Varchar |
| Status | Whether system is operations, in development, or decommissioned | Varchar |
| SDD_Owned | Indicates if the System is SDD Owned | Varchar |
| ISO | Designated FISMA system owner. | Varchar |
| ISSO | Information System Security Owner. Individual with assigned responsibility for maintaining the appropriate operational security posture for an information system or program. | Varchar |
| ISSO_Funding | Indicates who funds the ISSO | Varchar |
| ISSM | Information Systems Security Manager. Individual responsible for the information assurance of a program, organization, system, or enclave. | Varchar |
| Created | Indicates the time when the record was created | DateTime |
| Updated | Indicates the time when the record was updated | DateTime |
| **Moa_information_insight Table** | | |
| Key | Indicates the unique Key ID for this record | Varchar |
| MOA_Title | Indicates the MOA title | Varchar |
| MOA_ID | Indicates the MOA ID | Varchar |
| Status | Indicates the MOA Status | Varchar |
| Type | Indicates the MOA Type | Varchar |
| FY | Indicates the Fiscal Year the MOA is for | Varchar |
| MOA_Document | Indicates MOA document | Varchar |
| Comments | Field capturing any comments | Varchar |
| SDD_Customer | Indicates the SDD customer the MOA is for | Varchar |
| Customer_Program | Indicates the Customer Program the MOA is for | Varchar |
| Services | Indicates the Services provided in the MOA | Varchar |

| | | |
|---|---|---|
| Products | Indicates the Products provided in the MOA | Varchar |
| Updated | Indicates the time when the record was updated | DateTime |
| Created | Indicates the time when the record was created | DateTime |
| **Cfit Table** | | |
| FY | Indicates the Fiscal Year | Bigint |
| CFIT_ID | Unique CFIT ID for record | Bigint |
| MOAID | Indicates MOA ID | Double Precision |
| Office | Indicates the Office for the CFIT record | Text |
| Fund_Source | Indicates the Funding Source | Text |
| LOA | Indicates Line of Accounting (LOA) Code | Text |
| Fund_Code | Indicates the Funding Code | Text |
| WBS_Program | Indicates the Work Breakdown Structure (WBS) Program | Text |
| WBS_Project | Indicates the Work Breakdown Structure (WBS) Project | Text |
| Requirement_Description | Indicates the description for the CFIT record | Text |
| PlanComDt | Indicates the planned decommissioning date for record | Date |
| ComMo | Indicates the commissioned month | Text |
| Planned | Indicates when it was Planned | Double Precision |
| Committed | Indicates when it was Committed | Double Precision |
| Committed_Date | Indicates the Committed Date | Date |
| ComAge | Indicates the Commissioned Age in months | Bigint |
| Obligated | Indicates if it was obligated | Double Precision |
| Obligated_Date | Indicates the Obligation Date | Date |
| Expended | Indicates Expended amount | Double Precision |
| OE | Indicates Operational Expense (OE) | Double Precision |
| Remaining | Indicates remaining time on contract in months | Double Precision |
| PR_Number | Indicates the Procurement Request (PR) Number | Text |
| Qtr | Indicates the Quarter (Qtr) for the Record | Bigint |
| POP_Start | Indicates Period of Performance (POP) Start Date | Text |
| POP_End | Indicates Period of Performance (POP) End Date | Text |
| Task_Order_Contract# | Indicates Task Order/Contract # | Text |
| Vendor | Indicates Vendor Name | Text |
| COR | Indicate Contacting Officer Representative (COR) name | Text |
| CO | Indicates Contracting Officer (CO) name | Text |
| SubObject_Class | Indicate Subject Object Class | Text |
| Project_Code | Indicates Project Code | Text |

| | | |
|---|---|---|
| Task_Code | Indicates Task Code | Text |
| Program_Code | Indicates Program Code | Text |
| Organization_Code | Indicates Organization Code | Text |
| PPA | Indicates Post-Payment Audit (PPA) Code | Text |
| TAS Code | Indicates Treasury Account Symbol (TAS) Code | Text |
| Collected_Amount | Indicates the Collected Amount $ | Double Precision |
| Customer_LOA | Indicates the Customer Letter of Assessment | Text |
| Mission_Priority | Indicates Customer Mission Priority | Text |
| UII_Code | Indicates Unique Investment Identifier (UII) Code | Text |
| IAA_CCF | Indicates Inter-/Intra-agency Agreements IAA_CCF Code | Text |
| In_External | Indicates if the customer is Internal or External to SDD | Text |
| RefFSID | Indicates customer Ref Financial Services ID | Double Precision |
| ITAR_Number | Indicates Information Technology Acquisition Review ITAR Number | Text |
| Budget_ObjectClass | Indicates Budget Object Class | Text |
| CJ | Indicates Congressional Justification (CJ) | Text |
| ITAR_Status | Indicates Information Technology Acquisition Review (ITAR) Status | Text |
| Customer_POC | Indicates Customer Point of Contact (POC) | Text |
| DHS_Investment | Indicates DHS Investment | Text |
| Customer | Indicates Customer | Text |
| Contract_Number | Indicates the Contract Number | Text |
| Updated_by | Indicates who updated the record | Text |
| Updated | Indicates when the record was updated | Timestamp |
| ORG | Indicates the Organization | Text |
| Notes | Indicates Notes | Text |

### Data Retention Requirements

Per General Records Schedule 5.2 Item 020, “there is not a defined Records Retention Schedule and is the responsibility of the project team to determine based on their business needs”. Neptune and PPM have determined that data will be retained for seven (7) years. When the data retention date reaches 30 days prior to 7 years after the end of the calendar year it was created, DS shall contact PPM to determine if there is an ongoing business need to retain the data and if no business need exists, delete the data.

### Data Refresh Rates within Data Services

PPM data will be refreshed in the Neptune database bi-weekly (every two weeks) from.

## Law Enforcement Activity Reporting

**Components** Office of the Chief Information Officer (OCIO)

**Status** Approved

**Description**

The purpose of these reports is to provide information on Law Enforcement incidents across DHS. DSB is developing a database in Neptune to maintain this data. In that database, each report will be housed in its own schema and from those schemas, views will be built in the database. From those views, Tableau reports will be generated to provide an automated reporting mechanism for The Office of Strategy, Policy, and Plans to leverage.

**Relevant Compliance Documents**

PIA

DHS/ALL/PIA-046 DHS Data Framework

**Individuals Covered**

No PII is included in these datasets. The information included in these datasets is metadata associated with the incidents themselves, like the number of incidents, date of incident, reporting component, etc.

**Data Elements Covered**

| **column_name** | **field_description** | **target_data_type** |
|---|---|---|
| component_office | DHS Component or office reporting on this form | VARCHAR |
| fiscal_year | The period that this report is in reference to | DATE |

| | | |
|---|---|---|
| category | The category the type of reporting question falls under | VARCHAR |
| reporting_questions | The Lautenberg Reporting questions | INT |
| Sub-Component_Office | [Sub-Component/Office A] | INT |

### Data Retention Requirements

General Records Schedule 4.1 Item 020 “Records management program records” requires records to be destroyed no sooner than 6 years after the project, activity, or transaction is completed or superseded, but longer retention is authorized if needed for business use. As soon as the project support is no longer needed or superseded, the Data Services team will manually delete the records within 6 years of that date unless requested otherwise by the business.

### Data Refresh Rates within Data Services

The current refresh rate of the dashboard is daily.

## DHS Chief Human Capital Officer (CHCO) / Learning Technology and Innovation (LTI) / Headquarters Interim Training Site (HITS)

**Components** DHS Chief Human Capital Officer (CHCO)

**Status** Approved

**Description**

CHCO's Learning Technology and Innovation's (LTI) program manages the Headquarters Interim Training Site (HITS).HITS is an application in SharePoint supported by OCIO's online analytical processing platform called Neptune. HITS allows employees to complete mandatory training and to access their training completion status (name, course title, completed date). HITS will also generate reports that are exported to Neptune, fused with non-SPII employee attributes, and made accessible to EIE for Tableau reports that will allow authorized reviewers to see mandatory training completion based on the reviewer permissions. The HITS platform is designed for LTI to see all the learning and course trainings being assigned and completed by individuals throughout DHS.

**Relevant Compliance Documents**

PIA

DHS/ALL/PIA-046 DHS Data Framework[46]

DHS/ALL/PIA-049(a) Performance and Learning Management System (PALMS)[47]

SORN

DHS/All-003 Department of Homeland Security General Training Records

DHS/ALL-004 General Information Technology Access Account Records System (GITAARS)[48]

OPM/GOVT-1 General Personnel Records[49]

[46] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR DHS DATA FRAMEWORK DHS/ALL/PIA-046 DHS Data Framework (2020), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[47] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR Performance and Learning Management System (PALMS) DHS/ALL/PIA-049(a) (2017), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs..

[48] *See* DHS/ALL-004 General Information Technology Access Account Records System (GITAARS), 77 Fed. Reg. 70792 (November 27, 2012), *available at* https://www.dhs.gov/system-records-notices-sorns.

[49] *See* OPM/GOVT-1 General Personnel Records, 80 Fed. Reg. 74815 (November 30, 2015), *available at* https://www.dhs.gov/system-records-notices-sorns.

### Individuals Covered

PII (e.g., name, email) will be utilized and stored for DHS Employees/Contractors for trainings assigned, completed, and start and end dates. **Data Elements Covered**

The data elements below and their corresponding descriptions will be the datasets Neptune will source and store. The datasets will then be combined to create Tableau reports and dashboards to give insights on Trainings Assigned, Users, Completion Dates, Organization Codes, and Supervisors.

| Field Name | Description | Data Type |
|---|---|---|
| **Trusted Identity Exchange (TIE) Table** | | |
| Email address | Indicates the individuals DHS email address | Varchar |
| Organization code | Indicates the lowest level in the organizational hierarchy to which the person is assigned | Varchar |
| Person handle | Indicates the Unique system identifier for a Person generated by the ISMS system | Varchar |
| **SharePoint Table** | | |
| course_id | Indicates the unique course id of the training course on the HITS site | Varchar |
| training_title | Indicate the title of the training course on the HITS site | Varchar |
| start_date | Indicates when the user started the training | Date |
| end_date | Indicates when the user completed the training | Date |
| period_frequency | Indicates the number of days the training is valid i.e. 365,730, 0 = no expiration | Int |

### Data Retention Requirements

General Records Schedule 5.2 Item 020, “Records management program records” requires records to be destroyed no longer needed for business use, or according to an agency predetermined time period or business rule. When the data retention date reaches 30 days prior to 3 years after the end of the calendar year it was created, DSB shall contact HITS to determine if there is an ongoing business need to retain the data and if no business need exists, will then delete the data.

### Data Refresh Rates within Data Services

The current refresh rate of the dashboard is daily.

## Data Services Branch (DSB) - Office of the Chief Procurement Officer (OCPO) Enterprise Reporting Application (ERA)

**Components** Office of the Chief Procurement Officer (OCPO)

**Status** Approved

### Description

ERA is a tool to assist OCPO in performance and compliance management as well as organizational assessment of DHS procurement activities. ERA consists of multiple applications and tools used by the contracting workforce to execute, monitor, and report on the contracting process. Where data exists, ERA imports it and where data does not exist it is captured within ERA's applications and forms. To ensure that the ERA system and program are operating as efficiently and effectively as possible, DSB is tasked with designing and implementing independent validation and verifications services such as interviews to stakeholders. DSB will provide a comprehensive and thorough independent validation, verification, and assessment of ERA. These interviews will result in overall system and program management improvements.

### Relevant Compliance Documents

PIA

DHS/ALL/PIA-046 DHS Data Framework[50]

SORN

N/A

### Individuals Covered

ERA operators and administrators.

### Data Elements Covered

The project will conduct a comprehensive assessment of ERA system by interviewing stakeholders. Listed below are some field names in ERA system.

[50] *See* U.S. Department of Homeland Security, Privacy Impact Assessment for DHS Data Framework DHS/ALL/PIA-046 DHS Data Framework (2020), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

| Field Name | Description |
|---|---|
| PDS Category | This field indicates the general label name associated with a group of Procurement Data Standard (PDS) fields. |
| PDS Category Description | This field indicates brief details of the PDS Category field. |
| PDS Reference Number | This field indicates the unique number (Ref.) associated with this field in the PDS. |
| PDS Field Name | This field indicates the field name in the Procurement Data Standard. |
| Source System | This field indicates the source system or system of origin of the field. |
| FPDS Field Name | This field indicates the field name in Federal Procurement Data System (FPDS). |
| SAM Field Name | This field indicates the field name in System for Award Management (SAM). |
| PRISM Field Name | This field indicates the field name in Enterprise PRISM Instance (PRISM). |
| CLM Field Name | This field indicates the field name in Contracts Lifecycle Management (CLM). Currently listed as To Be Determined (TBD) until extracts are created. |
| CPARS Field Name | This field indicates the field name in the Contractor Performance Assessment Reporting System (CPARS). |
| Character Length | This field indicates the allowable numerical limit of characters (letters, numbers, etc.) |
| Character Type | This field indicates the kind of character data string. |
| XML Tag Name | This field indicates the XML Tag Name used for further reporting and programming. |
| Contract Phases | This field indicates the contract phases that apply to this field. |
| PDS Field Description | This field indicates brief details of the PDS field. |
| Expected Component Data | This field indicates details on data the component should provide regarding this field. |
| Inclusion Status | This field indicates whether this data is required or optional. |

### Data Retention Requirements

General Records Schedule 5.2 Item 020, “there is not a defined Records Retention Schedule and is the responsibility of the project team to determine based on their business needs.” Neptune and OCPO ERA have determined that data will be retained for seven (7) years.

### Data Refresh Rates within Data Services

There will be no data refreshes as all information will be obtained through a one-time set of interview assessments.

## Data Services Branch (DSB) - DHS Office of the Chief Information Officer/CBP Analytics

**Components** U.S. Customs and Border Protection

**Status** Approved

**Description**

The DSB will source CBP Analytics data from the CBP SharePoint site. The CBP Analytics program information is used to collect data regarding individuals using the CBP One mobile application. The CBP One mobile application is used to provide the public with a single portable access to a variety of CBP services such as appointment requests. CBP Analytics will utilize Neptune to source statistics and metrics for CBP One use. The DSB will provide subject matter expertise, analytics, engineering, and implementation services to improve data sourcing and data visualization. CBP Analytics will track the fraud, waste, and abuse detection in the CBP One opt-in process. This will help to design an automated workflow that can be used to provide trend analysis and provide multi-level analysis to enable key decision-making.

**Relevant Compliance Documents**

PIA

DHS/ALL/PIA-046 DHS Data Framework[51]

DHS/ALL/PIA-059 DHS Employee Collaboration Tools[52]

Forthcoming CBP SharePoint PIA - The data originally input in the CBP SharePoint site is covered by source system PIAs

SORN

DHS/ALL-004 General Information Technology Access Account Records System (GITAARS)[53]

The CBP data is covered by the source system SORNs for the data input into CBP SharePoint

[51] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR DHS DATA FRAMEWORK DHS/ALL/PIA-046 DHS Data Framework (2020), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[52] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR DHS EMPLOYEE COLLABORATION TOOLS DHS/ALL/PIA-059 (2017), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[53] *See* DHS/ALL-004 General Information Technology Access Account Records System (GITAARS), 77 Fed. Reg. 70792 (November 27, 2012), *available at* https://www.dhs.gov/system-records-notices-sorns.

### Individuals Covered

PII data (name, phone, email, date of birth, gender, and age) will be managed by CBP Analytics and will be utilized for the public and DHS Employees/Contractors who are users of the CBP One application.

### Data Elements Covered

The information listed below is used to collect data regarding users of CBP One application in order to assess the level of fraud associated with CBP One application use by individuals.

| Field Name | Description | Type |
|---|---|---|
| port_name | Port of entry location name | VARCHAR |
| confirmation_number | Indicates confirmation number | BIGINT |
| citizenship_code | Code to identify citizenship | VARCHAR |
| last_name | Indicates individual's last name | VARCHAR |
| first_name | Indicates individual's first name | VARCHAR |
| contact_email | Indicates applicant's contact email address. | VARCHAR |
| date_of_birth | Indicates applicant's date of birth | DATE |
| groupid | Indicates group ID | BIGINT |
| registration_date | Indicates appointment registration date | DATE |
| opt_in_date | Date when individual accessed the app to apply for an appointment | DATE |
| file_date | Shows date application was filed | DATE |
| confirmation_number | Shows submission or appointment confirmation number | BIGINT |
| first_name | Indicates individuals first name | VARCHAR |
| last_name | Indicates individual’s last name | VARCHAR |
| dob | Date of Birth | DATE |
| group_id | Identifies group of individuals that are registered together on | BIGINT |
| error | Indicates error file | VARCHAR |
| time_occured | Indicates occurrence time | DATE |

| | | |
|---|---|---|
| A-Number | Indicates a non-immigrant's alien or admission number | VARCHAR |
| Subject ID | Indicates identification for the subject | VARCHAR |
| FIN # | Fingerprint Identification Number | VARCHAR |
| CIV ID | Customer Identity Verification | VARCHAR |
| Event # | Indicates event number | VARCHAR |
| SIGMA Event # | Indicates SIGMA event # | VARCHAR |
| Office | Indicates office | VARCHAR |
| Arrest Sector / Field Office | Any one of 20 geographic areas into which the United States is divided for OFO activities, plus CBP preclearance activities abroad. | VARCHAR |
| Arrest Station / POE | Point of Entry | VARCHAR |
| Title Authority | The Title within the U.S. Code that provides the legal authority to process noncitizens for exclusion or removal. See Title 8 and Title 42. | VARCHAR |
| Arrest Date / Time | Date of detention for an individual by legal authority based on an alleged violation of the law. | DATE/TIME |
| Custody Status | Status of noncitizens apprehended by CBP and detained in CBP short-term custody facilities or processing centers (up to 72 hours) for purposes of case processing and disposition, pending transfer to another agency, release, or additional or other actions. | VARCHAR |
| Demographic | a particular sector of a population | VARCHAR |
| Age | The number of whole years from a person’s date of birth to an event. | VARCHAR |
| Date of Birth | Date which a person was born | DATE |
| Gender | Gender is a broad term typically associated with a person's own sense of their behaviors, characteristics, and thoughts, often in relation to their sex. | VARCHAR |
| Citizenship | Country to which a person owes allegiance | VARCHAR |

| | | |
|---|---|---|
| | and by which they are entitled to be protected. Each country sets its own rules for granting citizenship, which may be based on birth or naturalization, and some persons may not have a country of citizenship | |
| Separated or Disbanded Family? | Condition where family members are involuntarily separated from each other, often because of immigration systems | VARCHAR |
| Separated / Disbanded Reason | Reason why family members are involuntarily separated from each other, | VARCHAR |
| Family Makeup | Indicates related constituents of a family | VARCHAR |
| Family Size | The total number of people related by blood (consanguinity), marriage (affinity), or adoption that live together | NUM |
| Age / Gender Group | The number of whole years from a person's date of birth to an event. | VARCHAR |
| Medical Status | An assessment of a patient's current medical condition including medication | VARCHAR |
| General Processing Code (USBP Only) | Indicates general processing code for United States Border Patrol | VARCHAR |
| Custodial Redetermination | Judges authority to review or modify a person's (immigrant) custody status. | VARCHAR |
| Claimed Fear | A credible fear of persecution or torture is a "significant possibility" that you can establish in an Asylum Merits Interview before an asylum officer or in proceedings before an IJ, that you have been persecuted or have a well-founded fear on account of your race, religion, nationality, membership in a particular social group, or political opinion if returned to your country. | VARCHAR |
| Operational Code | Indicates operational code | VARCHAR |
| Manifest ID | Indicates manifest ID | VARCHAR |
| Manifest Location | Indicates manifest location | VARCHAR |
| Manifest Transfer Date / Time | Indicates manifest time | DATE |
| Manifest Notes | Shows manifest notes | VARCHAR |

| | | |
|---|---|---|
| Has CBP One Confirmation #? | Indicates if individual has a confirmation number from the CBP One app | VARCHAR |
| Intended Address Street 1 | Intended U.S physical address for individual entering the U.S | VARCHAR |
| Intended Address Street 2 | Intended U.S physical address for individual entering the U.S | VARCHAR |
| Intended Address Street 3 | Intended U.S physical address for individual entering the U.S | VARCHAR |
| Intended Address City | City where individual intends to reside upon entering the U.S | VARCHAR |
| Intended Address State | State where individual intends to reside upon entering the U.S | VARCHAR |
| Intended Address Zip Code | Zip code where individual intends to reside upon entering the U.S | VARCHAR |
| Intended Address Creation Date / Time | Date and time when intended address was entered in the CBP One app. | Date |
| Foreign Address State of Residence Code | Non-U.S code for state of residency | VARCHAR |
| Foreign Address State of Residence | Non-U.S state of residency | VARCHAR |
| Foreign Address Country of Residence | Non-U.S country of residency | VARCHAR |
| Border | Crossing area between the United States and foreign contiguous territory | VARCHAR |
| Mode | Method of travel (air, land, or sea) for persons arriving at ports of entry (POEs). | VARCHAR |
| port_code | Port of entry location code | VARCHAR |

### Data Retention Requirements

General Records Schedule 5.2 Item 020, “there is not a defined Records Retention Schedule and is the responsibility of the project team to determine based on their business needs.” Neptune and CBP Analytics have determined that data will be retained for seven (7) years.

### Data Refresh Rates within Data Services

There is no data refresh schedule. Data refresh is only done when there is new data provided and made available in SharePoint site.

## Data Services Branch (DSB) -The Office for Civil Rights and Civil Liberties (CRCL)/ServiceNow: Complaint Management System (CMS) 2.0

|  |  |
|---|---|
| **Components** | Civil Rights and Civil Liberties |
| **Status** | Approved |

**Description**

The CRCL project is designed for CRCL to see all the work being performed on the allegations coming from individuals and communities that might have been affected by DHS activities and policies. CRCL will utilize Neptune to source tasks, statistics, metrics, and progress for allegations of civil rights and civil liberties violations. This will help to design an automated workflow that can be used to improve data efficiency, provide trend analysis, and provide multi-level analysis to enable key decision-making. The information will be sourced from CRCL's ServiceNow CMS 2.0 system. This system includes data regarding individuals and communities whose civil rights and civil liberties may be affected by DHS activities.

**Relevant Compliance Documents**

PIA

DHS/ALL/PIA-046 DHS Data Framework[54]

DHS/ALL/PIA-029 Entellitrak[55]

DHS/ALL/PIA-059 DHS Employee Collaboration Tools[56]

SORN

DHS/ALL-029 Civil Rights and Civil Liberties Records[57]

**Individuals Covered**

To minimize PII overall, CRCL will not be taking any data from the Consumer and Constituent Profile tables in CMS 2.0, where names, A-Number, and other PII reside. Rather, CRCL will

[54] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR DHS DATA FRAMEWORK DHS/ALL/PIA-046 DHS Data Framework (2020), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[55] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR ENTELLITRAK DHS/ALL/PIA-029 (2010), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[56] *See* U.S. DEPARTMENT OF HOMELAND SECURITY, PRIVACY IMPACT ASSESSMENT FOR DHS EMPLOYEE COLLABORATION TOOLS DHS/ALL/PIA-059 (2017), *available at* https://www.dhs.gov/privacy-documents-department-wide-programs.

[57] *See* DHS/ALL-029 Civil Rights and Civil Liberties Records, 75 Fed. Reg. 39266 (July 8, 2010), *available at* https://www.dhs.gov/system-records-notices-sorns.

transfer data points about complaints that are disconnected from PII. Aside from the Summary of Allegation field, which usually contains names and A-Numbers, PII will largely be absent form the CRCL data inputted into Neptune. The Summary of Allegation field is used in dashboards to manage caseload, and thus important to CRCL's reporting needs.

### Data Elements Covered

The data elements below and their corresponding descriptions will be the datasets Neptune will source and store. The datasets will then be combined to create Tableau reports and dashboards to give insights on Allegations Opened/Closed, Component Assigned to Review, Workflow, and Incident Locations. The below data fields are a sample of the larger data sets. Neptune will provide full listing of tables and fields upon request.

| Field Name | Description | Data Type |
|---|---|---|
| **sn_gsm_crcl_allegation (sn_gsm_crcl_allegation) Table** | | |
| u_504_complaint_determination | Indicates the CRCL Allegation Section 504 Complaint Determination | text |
| u_medical_referral_decision_comments | Indicates the CRCL Allegation Medical Referral Comments | text |
| watch_list | Indicates the CRCL Allegation is on Watch List | text |
| caused_by | Indicates the CRCL Allegation of who may have Caused the Issue | text |
| **metric_definition (metric_definition) Table** | | |
| description | Indicates the Metric Description | Text |
| sys_mod_count | Indicates the number of times the Record Issue was Modified | Text |

### Data Retention Requirements

Disposition: N1-563-07-6-2

All related and relevant documents, and the component's memorandum of resolution in accordance with records schedule N1-563-07-6, b.1 and will destroy or delete seven years after resolution or closure of the case.

Retained issues are either maintained by CRCL because of the significance of the issue, which may result in policy change, or issues retuned from the component for resolution in accordance with N1-563-07-6, b.2 and will destroy or delete 75 years after resolution or closure of the case.

Significant case files involve allegations made against senior DHS officials; attract national media or congressional attention; present significant or novel questions of law or policy; and result in substantive changes in DHS policies and procedures. Significant case files will be selected by the Headquarters and component civil rights and civil liberties offices based on these criteria. In accordance with N1-563-07-6, b.3 records are maintained through the end of fiscal year in which the significant case file is closed. Records are transferred to NARA five years after the case is closed according to NARA transfer guidance and regulations.

The system will purge records in accordance with the appropriate retention schedule and delete automatically.

### Data Refresh Rates within Data Services

The current refresh rate of the dashboard is hourly.

# Appendix B: Approved Mission Uses

*Last updated: November 2, 2022*

Appendix B includes details and information on approved uses for which data in Neptune can be used by DHS components. If the approved purposes or access areas change for the Data Services Branch, this Appendix will be updated accordingly.

The various capabilities delivered by the Data Services Branch will provide value across many DHS mission areas. Authorization to use Neptune is granted by evaluating the authorities and policies relating to that mission as individuals or systems access the data in the Data Services Branch.

1. **Benefits and Authorizations** — Benefits are programs, projects, services, and activities provided by DHS that directly assist individuals or groups of individuals. Authorizations are grants of permission to engage in specified activities that are proscribed by law or otherwise regulated. The Data Services Branch has been approved to assist with Immigration benefits, specifically the conferral, certification, change, adjustment, or extension of any status granted under the Immigration and Nationality Act.[58]
2. **Law Enforcement** — Activities directed toward the preservation of public order and safety, including protection of persons and property (real and other) in accordance with a statutory authority. The Data Services Branch has not added any users for law enforcement purposes.
3. **National Security** — The comprehensive program of integrated policies and procedures for the departments, agencies, and functions of the United States Government aimed at protecting the territory, population, infrastructure, institutions, values, and global interests of the Nation. The Data Services Branch has been approved for the following national security mission uses:
   a) **Border Security:** The protection of U.S. borders from the illegal movement of weapons, drugs, contraband, and people, while promoting lawful entry and exit,

[58] 8 U.S.C. § 1572.

to include the disruption and dismantling of transnational organizations that engage in smuggling and trafficking across the U.S. border.[59]

b) **Foreign Nationals:** The care, protection, and tracking of asylum seekers, refugees, and other foreign nationals OCONUS and CONUS.

4. **DHS Employee Efficiencies** – Efficiencies are programs, projects, services, and activities provided by DHS that directly assist individuals in performing their functions and in meeting the mission of DHS by assisting in the ability make more informed strategic and policy decisions. DHS employees is meant to include DHS employees, as well as federal contractors.

5. **Health Information:** Information regarding health status of foreign nationals or DHS staff as it is required for strategic decisions or in relation to placement or care decisions.

[59] 6 U.S.C. § 202.

# Appendix C: Approved Users

*Last updated: November 2, 2022*

Appendix C includes details and information on the authorized users of the Data Services Branch. If the list of authorized users changes for the Data Services Branch, this Appendix will be updated accordingly.

Users are granted access to the Data Services Branch provided they are authorized to access the data for a purpose detailed in Appendix B. The following list identifies the users authorized to utilize the Data Services Branch, by Component and Program. Each use of the data is documented in a detailed Mission Use Case approved by the Data Framework Working Group (DFWG) made up of the DHS Privacy Office (PRIV), the Office of Civil Rights and Civil Liberties (CRCL), and the Office of General Counsel (OGC). The approved Mission Use Cases are summarized below.

## 1. DHS Office of Immigration Statistics (OIS)

| | |
|---|---|
| **Organization(s):** | Office of Immigration Statistics (OIS) |
| **Date Approved:** | January 30, 2020 |
| **Authorized Purpose:** | Benefits Adjudication – Immigration Benefits |

**High-Level Mission Use Case:**

The U.S. Department of Homeland Security's Office of Immigration Statistics (OIS) is responsible for leading the collection and dissemination of statistical information and analysis on the impact of immigration laws, migration flows, and immigration enforcement to Congress and the public. OIS's goal is to provide high quality statistical information that is relevant, timely, cost-effective and customer-oriented. OIS provides reporting through The Yearbook of Immigration Statistics and annual OIS Statistical Data Production and Reporting (SDPR) reports and population estimates, which provide details of lawful permanent residents, refugees and asylees, naturalizations, nonimmigrant admissions, and enforcement actions. OIS also publishes additional reports at the request of the White House, Congress, and DHS' Secretary on an as-needed basis.

## 2. DHS Operation Vaccinate Our Workforce

| | |
|---|---|
| **Organization(s):** | Office of the Chief Information Officer (OCIO) |
| **Date Approved:** | April 28, 2020 |
| **Authorized Purpose:** | DHS Employee Efficiencies – Covid-19 Vaccination Effort |

**High-Level Mission Use Case:**

The Operation Vaccinate Our Workforce (VOW) was created to protect DHS frontline workforce and the public with whom they interact. Through Operation VOW, DHS is pursuing new and existing ways to identify and prioritize COVID-19 vaccine allocations for frontline personnel for whom teleworking is not an option.

The Operation VOW response data will be used to create visual dashboards through business intelligence services like Tableau to report staff responses to the VOW initiative. The responses will be used to gauge a given DHS staff's desire to receive the COVID-19 vaccination, if they have already received a dose, or if they have an appointment to receive 1 of 2 doses or a single full dose.

3. **DHS Hummingbird on Neptune**

**Organization(s):** Office of the Chief Information Officer (OCIO)

**Date Approved:** November 2021

**Authorized Purpose:** National Security

**High-Level Mission Use Case:**

The Hummingbird application is a cloud-based case management system supported by ServiceNow on the Department of State (DoS) network. This system is a custom application built on the ServiceNow platform. The Data Lake system and processes are a storage repository that hold raw data from Hummingbird and other systems. These systems allow for the curation of data for reporting. At the present moment, we will maintain a continuous connection between the DOS ServiceNow platform and the DHS Data Lake. The data lake currently only connects to the Hummingbird application that is built on the DOS ServiceNow platform.