Privacy Impact Assessment Update
for the

DHS Data Framework – External Sharing

DHS/ALL/PIA-046(c)

March 30, 2016

Contact Point
Paul Reynolds
Data Framework Program Management Office
Department of Homeland Security
(202) 447-3000

Reviewing Official
Karen L. Neuman
Chief Privacy Officer
Department of Homeland Security
(202) 343-1717

Abstract

The DHS Data Framework is DHS's "big data" solution to build in privacy protections while enabling more controlled, effective, and efficient use of existing homeland security-related information. The DHS Data Framework includes the Neptune and Cerberus systems. DHS is updating the Data Framework Privacy Impact Assessment to reflect that DHS will now use Cerberus to share information externally, including "bulk information sharing," with U.S. Government partners, consistent with information sharing access agreements, published Privacy Impact Assessments, and Systems of Record Notices for the underlying source systems of the DHS Data Framework.

Introduction

In a Privacy Impact Assessment (PIA) published on November 6, 2013, and PIA updates published on August 29, 2014, and February 27, 2015, the Department of Homeland Security (DHS) previously described the Department's development of the DHS Data Framework ("Framework").[1] Currently, the Framework includes the Neptune[2] and Cerberus[3] systems.

Neptune is the unclassified "data lake" that DHS uses to receive, store, and tag the data from unclassified DHS information technology (IT) systems. Once tagged, unclassified DHS data sets from Neptune are transferred to Cerberus, which is the classified data lake that DHS uses to perform classified searches of unclassified DHS data sets.

Reason for the PIA Update

To date, only DHS users have been able to access data through the Framework. DHS will now use Cerberus to share information externally, including "bulk information sharing,"[4] with U.S.

---

[1] Please see the following privacy impact assessments: DHS/ALL/PIA-046 DHS Data Framework, November 6, 2013, DHS/ALL/PIA-046(a) DHS Data Framework, August 29, 2014 and DHS/ALL/PIA-046(b) DHS Data Framework, February 27, 2015.

[2] For more information about Neptune, please see the following privacy impact assessments: DHS/ALL/PIA-046-1 Neptune Pilot, September 25, 2013; DHS/ALL/PIA-046-1(a) Neptune Pilot, August 29, 2014; and DHS/ALL/PIA-046-1(b), February 27, 2015.

[3] For more information about Cerberus, please see the following privacy impact assessments: DHS/ALL/PIA-046-3 Cerberus Pilot, November 22, 2013; DHS/ALL/PIA-046-3(a) Cerberus Pilot, August 29, 2014; and DHS/ALL/PIA-046-3(b) Cerberus, February 27, 2015.

[4] Bulk information sharing is not linked to a particular number of records. Instead, bulk information sharing is determined by whether the information shared is sufficiently tailored to reasonably exclude information that is not relevant to a partner's request for data. Bulk information sharing refers to the transmission of large quantities of intelligence or information, which, due to technical or operational considerations, is transmitted without the use of discriminants reasonably likely to exclude any intelligence or information not relevant to the need giving rise to the

Government partners. Neptune will provide Cerberus with the Framework data that Cerberus shares externally, but Neptune itself will not share any information external to DHS. This overarching PIA update for the Data Framework covers external sharing from Cerberus, including using data provided by Neptune. This PIA provides notice that Cerberus may be used to share information with U.S government partners and DHS is not updating the respective information technology system PIAs for Cerberus and Neptune.

Bulk information sharing involves the transfer of all or portions of a Privacy Act System of Records (SORN) outside of DHS to another U.S. Government partner. For the reasons described in the preceding footnote, not all external sharing will qualify as bulk information sharing. For example, if DHS shares a narrowly tailored data sample (e.g., a list of known or suspected terrorists) with a partner to identify known or suspected terrorists, then the sample would not qualify as bulk information sharing. Whether the sharing is in bulk does not impact the oversight or controls DHS applies to external sharing that is leveraging the Framework, but the Department notes the distinction for transparency purposes.

Cerberus provides the technical architecture to reduce or replace existing external information sharing with U.S. Government partners. Additionally, Cerberus allows DHS to share information through secure, automated connections rather than delivering data through *ad hoc* transfers of portable media (e.g., discs or hard drives).

External sharing, including bulk sharing, may occur via the Cerberus architecture. Any external sharing activities that use the Cerberus architecture must be described in the appropriate PIA for those projects and covered by a routine use in the SORN for the source DHS IT system. Additionally, all external sharing through Cerberus must be governed by an Information Sharing and Access Agreement (ISAA), such as a Memorandum of Agreement (MOA).

No external users are able to access data or tools within the Cerberus system, although DHS anticipates developing such a capability in future iterations of the Framework, when appropriate technical and policy controls have been implemented. Similarly, Neptune does not share information externally, although Neptune, like Cerberus, may be accessed by external users in future iterations of the Framework.

Fair Information Practice Principles (FIPPs)

The Department applies the following Fair Information Practice Principles, developed from

recipient's request (specific identifiers, selection terms, etc.). For example, transmitting a list of known or suspected terrorists in response to a request for such information would not constitute a bulk data transfer because the request and its response are limited in scope to information reasonably likely to be of value to the recipient, but transmitting information about a group of individuals for the purpose of identifying known or suspected terrorists within that group would qualify as a bulk data transfer.

the Privacy Act's underlying concepts, to account for the nature and purpose of the information being collected in relation to the Department's missions. While some of the principles analysis remains unchanged from the initial Framework phases, the privacy impacts resulting from sharing information outside of the Department via Cerberus require additional analysis.

As described above, the creation of a robust governance structure is a principal means through which the Department intends to enhance the Framework's adherence to the Fair Information Practice Principles and further ensure the proper privacy, civil rights, and civil liberties protections are in place for the Framework.

1. Principle of Transparency

Principle: DHS should be transparent and provide notice to the individual regarding its collection, use, dissemination, and maintenance of personally identifiable information (PII). Technologies or systems using PII must be described in a SORN and PIA, as appropriate. There should be no system the existence of which is a secret.

Privacy Risk: There is a risk that individuals may not be aware their PII is being shared with external partners.

Mitigation: Any external sharing activities involving the Cerberus architecture must be described in the appropriate PIA for those projects and be covered by a routine use in the SORN for the system from which the DHS data originates. Prior to external sharing via Cerberus, the DHS Privacy Office will conduct a specialized Information Sharing Access Privacy Threshold Analyses (ISA-PTA), which is the Department's internal process to assess whether the applicable PIA and SORN provides appropriate notice of an information sharing activity or whether changes to the PIA and SORN are required.

There may be instances where national security concerns prevent DHS from describing an external sharing relationship in detail in a public PIA. In these instances, the sharing will be documented very generally in a public PIA with a classified PIA to provide a more in-depth assessment of the privacy risks and mitigations. While classified PIAs are not accessible to the public, they do document the privacy protections of a program and therefore provide internal oversight of DHS activities. Classified PIAs are available to appropriately cleared members of internal (e.g., Office of the Inspector General) and external (e.g., Privacy and Civil Liberties Oversight Board) oversight organizations.

Finally, to ensure appropriate institutional knowledge and to facilitate oversight, the DHS Privacy Office will use the ISA-PTA to document approved external sharing in a non-public appendix to this PIA that will be kept on file with the DHS Privacy Office.

2. Principle of Individual Participation

Principle: DHS should involve the individual in the process of using PII. DHS should, to the extent practical, seek individual consent for the collection, use, dissemination, and maintenance of PII and should provide mechanisms for appropriate access, correction, and redress regarding DHS's use of PII.

Privacy Risk: There is a risk that an individual will not be able to receive appropriate access, correction, and redress regarding the use of his or her PII because the information resides outside of a DHS system.

Mitigation: Once DHS releases information pursuant to a Routine Use under the Privacy Act, the recipient agency or department may apply its own access, correction, and redress procedures, which may or may not be similar to DHS's. To partially mitigate this risk, DHS documents the access, correction, and redress procedures for DHS data in an ISAA with the external partner. The agreement may, as a matter of DHS policy, stipulate additional access, correction, or redress procedures for DHS data shared with an external partner.

The DHS Data Access Request Council (DARC) is the coordinated oversight and compliance mechanism that reviews bulk transfers of data in support of DHS's national or homeland security missions. The DARC ensures bulk sharing initiatives or activities comply with applicable law and adequately protect the privacy, civil rights, and civil liberties of the individuals whose information is shared. The DARC includes representatives from the DHS Privacy Office, Office for Civil Rights and Civil Liberties, and the Office of the General Counsel, collectively known as "DHS Oversight Offices." The DARC also includes representatives from DHS Components, including Component privacy officials and mission representatives. Part of the DARC's review of sharing agreements includes assessing whether appropriate access, correction, and redress procedures exist, and these procedures are documented in the ISAA.

3. Principle of Purpose Specification

Principle: DHS should specifically articulate the authority that permits the collection of PII and specifically articulate the purpose or purposes for which the PII is intended to be used.

Privacy Risk: There is a risk that DHS will use the Framework to share data with external partners for a purpose other than the purpose for which is was collected in the original DHS IT system.

Mitigation: This risk is partially mitigated. First, all external sharing must be documented in the appropriate source system PIA and SORN. For all external sharing through Cerberus, the DHS Privacy Office will conduct Privacy Threshold Analyses, which is the Department's internal process to assess whether the applicable PIAs and SORNs provide appropriate notice or if changes to the PIA and SORN are required. This process includes coordinating with the Component privacy offices.

Second, through the DHS DARC, the DHS Privacy Office performs a review of the sharing to ensure it is compatible with the purpose for which it was collected. Third, DHS documents in the ISAA the Privacy Act Routine Uses under which DHS is releasing the information and the purposes for which the recipient may use the data.

4. Principle of Data Minimization

Principle: DHS should only collect PII that is directly relevant and necessary to accomplish the specified purpose(s) and only retain PII for as long as is necessary to fulfill the specified purpose(s). PII should be disposed of in accordance with DHS records disposition schedules as approved by the National Archives and Records Administration (NARA).

Privacy Risk: There is a risk that permitting external sharing through the Framework will encourage DHS to replicate data sets outside of the Department.

Mitigation: This risk is not mitigated during this phase of the Framework. However, this risk is not new or specific to the Framework. Permitting external sharing through Cerberus may make it easier for DHS to share information with external partners and therefore generate an increase in sharing arrangements. DHS already has multiple external sharing relationships with external partners. However, DHS shares information as a method of last resort and only when mission needs necessitate the sharing. These external sharing arrangements occur with or without the Framework. By using Cerberus for data transfers, DHS can: (1) ensure the data is transferred through a secure mechanism and reduce dependence on portable media; (2) more easily refresh data, to promote data quality; and (3) provide robust oversight through the DARC and the Data Framework governance process. Finally, an important long-term goal of the Framework is to reduce or replace the number of external transfers of DHS data. In the long-term DHS plans to allow external partners to access DHS data within the Framework, which should eliminate the need for some transfers of DHS data.

Privacy Risk: There is a risk that data will be retained in an external partner's system for longer than is permitted.

Mitigation: This risk is partially mitigated through three mechanisms. First, sharing through Cerberus allows DHS to more easily refresh data with external partners, which also allows DHS to more easily update data that has been changed, deleted, or updated in source IT system to ensure that the appropriate retention periods are followed. This update process is an improvement over the use of other methods, such as the exchange of portable media, which provide less reliable means for enforcing retention periods through data updates. Second, for some projects, Cerberus allows DHS to share data that is tagged with the appropriate retention period, allowing DHS or a partner to perform deletions as needed. Third, DHS documents the appropriate retention period in the ISAA with the external partner. The DARC provides oversight for the ISAA and evaluates the partner's compliance with the agreement's terms, including those related to retention.

5. Principle of Use Limitation

Principle: DHS should use PII solely for the purpose(s) specified in the notice. Sharing PII outside the Department should be for a purpose compatible with the purpose for which the PII was collected.

Privacy Risk: There is a risk that once the data is transferred outside of the Department, external users will use the data for purposes other than those authorized.

Mitigation: This risk is partially mitigated. The DARC oversees bulk sharing with external partners and documents the authorized uses of DHS data in an ISAA with the external partner(s). DHS uses a variety of mechanisms to ensure compliance with ISAAs, including training for end users, periodic reviews of the agreement once implemented, briefings on the use of DHS data, and, if appropriate, a formal Privacy Compliance Review.

6. Principle of Data Quality and Integrity

Principle: DHS should, to the extent practical, ensure that PII is accurate, relevant, timely, and complete within the context of each use of the PII.

Privacy Risk: There is a risk that PII transferred outside of the Department will not be accurate, relevant, timely, or complete.

Mitigation: This risk is partially mitigated. Anytime DHS transfers data outside of a source system, there is a risk that changes will not be replicated to the external copy of the data. To help mitigate this risk, the ISAAs and PIAs that cover external sharing outline procedures to ensure data accuracy, such as data integrity checks and validation of results with DHS, as appropriate and applicable.

The use of Cerberus' technical architecture to transfer data in a secure, automated fashion will reduce this risk compared to external sharing efforts that rely on the exchange of portable media. Because of the automated sharing, DHS will be able to more easily (and in some cases more frequently) refresh data shared with external partners.

7. Principle of Security

Principle: DHS should protect PII (in all forms) through appropriate security safeguards against risks such as loss, unauthorized access or use, destruction, modification, or unintended or inappropriate disclosure.

Privacy Risk: There is a risk that data will be transferred to external partners in a manner that is insecure.

Mitigation: The Department follows the requirements for information assurance and security and the development of sensitive systems[5] and handling of sensitive information[6] for the Framework systems. Both Framework systems, Cerberus and Neptune, have system security plans and the Chief Information Security Officer's approval for Authority to Operate. Security and policy based controls enforced with these systems include:

- Data encryption of any media and during any transmission of data to prevent PII exposure.
- Only users with administrator privileges will be able to directly access the delivered data within Cerberus for data quality processing, to initiate external transfer for information sharing, or for other technical functions. These users will be vetted and approved for access up to the highest level of data on the system.

Finally, the use of Cerberus' automated, secure connection for external transfers represents an improvement in security compared to some of the existing external sharing arrangements, which use portable media.

Privacy Risk: There is a risk that external partners' systems do not adequately protect DHS data once it is delivered.

Mitigation: DHS's ISAAs with external partners require that reasonable physical, electronic, and procedural safeguards be maintained to appropriately protect the DHS information shared against loss, theft, misuse, and unauthorized access, disclosure, copying, use, modification, or deletion.

8. Principle of Accountability and Auditing

Principle: DHS should be accountable for complying with these principles, providing training to all employees and contractors who use PII, and should audit the actual use of PII to demonstrate compliance with these principles and all applicable privacy protection requirements.

Privacy Risk: There is a risk that the external partner will not comply with the ISAA.

Mitigation: DHS uses a variety of mechanisms to ensure compliance with ISAAs, including training for end users, periodic reviews of the agreement, briefings on the use of DHS data, and, if appropriate, a formal Privacy Compliance Review. DHS agreements also include provides to allow DHS to terminate an ISAA in the external party violates the agreement.

[5] *See* DHS 4300A Sensitive Systems Handbook.
[6] *See* DHS Handbook for Safeguarding Sensitive Personally Identifiable Information.

Conclusion

DHS developed the Framework specifically to ensure that it is consistently using DHS data for the purposes for which it was collected. There are several privacy risks to the overall Framework that have been mitigated, helping to demonstrate the ability to meaningfully use technology including dynamic access controls to mitigate risk. As the Framework continues to mature, this Privacy Impact Assessment will be updated periodically to account for any major changes to the information architecture and data governance.

Responsible Officials

Donna Roy
Office of Chief Information Officer
Department of Homeland Security

Michael Freeman
Office of Intelligence and Analysis
Department of Homeland Security

Approval Signature

Original signed copy on file with the DHS Privacy Office.

_____

Karen L. Neuman
Chief Privacy Officer
Department of Homeland Security