Privacy Impact Assessment
for the

Data Streaming Services

DHS/USCIS/PIA-078

March 20, 2019

Contact Point
Donald Hawkins
Privacy Officer
U.S. Citizenship and Immigration Services
(202) 272-8030

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

Abstract

U.S. Citizenship and Immigration Services (USCIS) uses Data Streaming Services as intermediary messengers to effectively and efficiently move data among USCIS systems in near real-time. The use of these services allows USCIS to transport data without the technical and administrative burden usually placed on the operating systems. USCIS is publishing this Privacy Impact Assessment (PIA) to evaluate the privacy risks and mitigations associated with the transport of personally identifiable information (PII) using these services.

Overview

The Department of Homeland Security (DHS), U.S. Citizenship and Immigration Services (USCIS) collects, maintains, uses, and disseminates large amounts of personally identifiable information (PII) related to administering and processing benefit requests for all immigrant and nonimmigrant benefits, as well as employee human resource data. USCIS identified the need to obtain a middleware platform to support the dissemination and storage of data between source systems and recipient systems to mitigate the impact to the source system's availability and operational functionality.

USCIS acquired Apache Kafka® (hereafter referred to as "Kafka") and various commercial middleware tools to effectively and efficiently move data among USCIS systems in near real-time. Kafka provides high availability and resiliency for data and uses an event-driven design when events or changes to a source system trigger an update to the recipient system. The implementation of Kafka and various commercial middleware tools, collectively referred to as "Data Streaming Services," minimizes the need for USCIS to customize protocols and communication methods to move data between USCIS systems, which compromises the accuracy of the data and often consists of full-time jobs for information technology staff that result in an increase of USCIS expenditures.

Data Streaming Services

The Data Streaming Services are a combination of data delivery tools and connections to facilitate the seamless communication between different USCIS systems. The Data Streaming Services currently includes Kafka and various commercial middleware tools. These services extract, replicate, and transport identified data sets from one system to another. The Data Streaming Services duplicates and shares a real-time mirror image of the information from the source system, providing reliable information in support of USCIS operations. This PIA evaluates the PII each data delivery tool collects and uses to operate Data Streaming Services and evaluates the privacy risks and mitigation strategies built into each data delivery tool and connection. USCIS plans to update this PIA as additional tools are added to support Data Streaming Services.

Kafka

Kafka is a distributed messaging system providing fast, highly scalable, and redundant messaging through a "publish and subscribe" model.[1] In this model, Kafka centralizes communication between producers[2] and consumers[3] of data. Kafka's distributed design gives it several advantages. Kafka is highly available and resilient to system failures and supports automatic recovery. These characteristics make Kafka an ideal fit for communication and integration between USCIS, DHS, and external systems.

All Kafka messages, or information (including PII), that are sent from the producer to a consumer, are organized into *topics*. A topic is a category of records that is published and stored in a message. Topics in Kafka can have many consumers that subscribe to the data.

To facilitate the effective and efficient transport of data between systems, Kafka has four core application programming interfaces (API):

- The **Producer API** allows a system to publish a stream of records to one or more Kafka topics.[4] For example, Kafka may retrieve identified topics from a source system (i.e., CLAIMS 3[5]) to make available and share with a consumer once subscribed to an identified topic. The information then resides in destination system (i.e., Customer Profile Management System (CPMS)[6]).
- The **Consumer API** allows a system to subscribe to one or more topics that are aligned to specific tables and data sets from the source system, and process the stream of records produced to them.[7] For example, a consumer (i.e., CPMS) subscribes to a source system (i.e., CLAIMS 3) to retrieve requested topics.
- The **Streams API** identifies amended data in the producer system to detect anomalies, fraud, or abnormal changes.[8] For example, a name change occurs in CLAIMS 3 and is automatically amended in CPMS.
- The **Connector API** allows building and running reusable producer or consumer systems that connect Kafka topics to existing systems.[9]

---

[1] In a publish and subscribe model, any message published to a topic is immediately received by all of the subscribers to the topic.
[2] A system that sends the messages. A producer push messages into a Kafka topic.
[3] A system that receives the messages. A consumer pulls messages off of a Kafka topic.
[4] The Producer API allows applications to send streams of data to topics in the Kafka cluster.
[5] *See* DHS/USCIS/PIA-016(a) Computer Linked Application Information Management System and Associated Systems, *available at* www.dhs.gov/privacy.
[6] *See* DHS/USCIS/PIA-060 Customer Profile Management System (CPMS), *available at* www.dhs.gov/privacy.
[7] The Consumer API allows applications to read streams of data from topics in the Kafka cluster.
[8] The Streams API allows transforming streams of data from input topics to output topics.
[9] The Connector API allows implementing connectors that continually pull from some source data system into

These APIs are the building blocks of Kafka and work harmoniously to extract, replicate, and load data from a producer to a consumer system.

Kafka is an intermediary back-end platform that does not have a user interface. A system connection is required to access a source system data set extracted by Kafka. No system gains access to this stream without undergoing the processes to (1) formally request access, (2) be provided with the subscription credentials/accounts, and (3) be configured to receive the specific streams subscribed from the originating system. Kafka relies on change data capture (CDC), a software design pattern to track the data that has changed. This approach ensures data integration based on the identification, capture, and delivery of changes made to data sources.

Middleware Tools

Kafka provides the transmission mechanism for these data feeds. Kafka requires a separate middleware layer to translate the information from the data sources before it enters Kafka. Once it exists Kafka the data interpreter translates the information into a readable format from the source system before it enters the destination repository. USCIS currently uses the various commercial middleware tools as the middleware layer with Kafka.

A comprehensive middleware tool is designed for real-time data integration and replication. The designated middleware tools read, extract, replicate, and load the data into the Kafka topics. It does not modify or perform any action that would impact or change the integrity of the data. It only extracts, replicates, and loads the data to Kafka topics, focusing solely on transportation. As changes are made to a USCIS source system the designated middleware tool reports and updates these changes to the topics identified in Kafka in real time. No changes are made to the source system data through Kafka or the designated middleware tool to ensure its integrity as the replicated data is transmitted to its destination repository. The middleware tool is a necessary requirement that facilitates access from the source and destination systems to Kafka.

USCIS Systems Using Data Streaming Services

USCIS is publishing this PIA to provide transparency to the overall use of data streaming services (i.e., Kafka and various commercial middleware tools). USCIS plans to implement Kafka by continuously establishing new producers, consumers, and topics of Kafka to facilitate effective and efficient transportation of USCIS data. As such, USCIS will frequently update the producer and consumer appendices to this PIA as new producers and consumers are respectively integrated into the Data Streaming Services.

Kafka or push from Kafka into a data system.

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?

Section 103 of the Immigration and Nationality Act (INA) authorizes USCIS to use Data Streaming Services to support the administration and adjudication of benefits.[10]

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

Data Streaming Services extract and replicate information from producer systems for consumer system use. Data Streaming Services rely on the source system SORNs to cover the collection, maintenance, and use of the source system data. The appendices to this PIA list the applicable SORN for each used by the Data Streaming Services.

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

Yes. Kafka and various middleware tools are minor applications under the Cloud Hosted Environment (CHE). The CHE Authority to Operate (ATO) is pending the publication of this PIA. CHE will enter into the Ongoing Authorization program, upon completion of this PIA. Ongoing Authorization requires CHE to be reviewed on a monthly basis and to maintain its security and privacy posture in order to retain its ATO. CHE security controls and organizational risks are assessed and analyzed (that vary by security control) to support risk-based security decisions. CHE also undergoes regular security audits to assess CHEs security compliance.

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

USCIS retains the audit logs associated with these services in accordance with General Records Schedule DAA-GRS2013-0006-0003, which states that the records are destroyed when the business use ceases. USCIS identified that it has the need to maintain audit logs for seven days to ensure the successful transport of data from producer to consumer. Data Streaming Services does not retain any data; it is merely a pass-through.

[10] *See* 8 U.S.C. § 1103.

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 Data Streaming Services are not subject to PRA requirements because it only extracts, transports, and loads information from producer systems to consumer systems. However, information from the producers may be subject to the PRA. Please see the respective producer PIA for PRA applicability.

Section 2.0 Characterization of the Information

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

Data Streaming Services are used as middleware transporting tools from data originating from a producer system to a consumer system. Data Streaming Services extract information from the producer system and replicate the data into staging table. Data Streaming Services then loads the information into the consumer system. The information Kafka transports from a producer system to a consumer system may contain PII. This primarily consists of information relating to (1) persons who have filed, for themselves or on the behalf of others (benefit requestors and beneficiaries), benefit requests for immigration benefits under the Immigration and Nationality Act, as amended, and/or who have submitted fee payments or received refunds from such benefit requests; (2) current, former, and potential derivatives of benefit requestors (family members); (3) sponsors (e.g., employers, law enforcement officers, individuals); (4) attorneys and representatives recognized by USCIS and/or accredited by the Board of Immigration Appeals (Representatives); (5) interpreters; (6) individuals who assist in the preparation of the benefit request (Preparers); (7) individuals who make fee payments on behalf of the benefit requestor; and (8) physicians who conduct immigration related medical examinations. Kafka may also transport data on USCIS employees.

The information extracted, replicated, and loaded by the Data Streaming Services is determined by the respective topic. The information may include, but is not limited to the following categories of information: names, contact information, birth information, unique identifiers, citizenship/nationality information, immigration status, marital status/family information, physical characteristics, employment information, immigration file tracking location, and background check results.

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

USCIS Data Streaming Services are platforms and middleware designed to transport data from a producer system to a consumer system. See the appendices to this PIA for specific producers and consumers.

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

USCIS Data Streaming Services include Kafka and various commercial middleware tools. These are commercial tools used to extract, replicate, and load data from one location to another. Furthermore, some of the USCIS Data Streaming Services may collect, use, or maintain information obtained from commercial sources or publicly available information previously maintained in the producer system. See the appendices to this PIA that detail the USCIS systems that may collect and use commercial and publicly available data.

2.4 Discuss how accuracy of the data is ensured.

The Data Streaming Services does not collect data directly from individuals but rather obtains data from other USCIS source systems. The Data Streaming Services depend on the accuracy and quality of data from each source system. Kafka and middleware tools ensure the accuracy of the data by collecting the information directly from the source systems. USCIS routinely reviews the transactional logs to ensure that the data remains accurate and complete during the transmission of data from the producer system to the consumer system.

For data that may be transferred through the Data Streaming Services, the data is encrypted and transported from the producer systems and is delivered "as is" to the consumer system with the exception of reformatting to standardize the representation of the data using the designated middleware tool. This process ensures the data integrity from the producer system to the consumer system. Any checks for accuracy of the data are accomplished at the producer system, and are out of scope of the Data Streaming Services and controls. The Data Streaming Systems cannot and does not provide any assurance that the data it delivers is accurate.

These Data Streaming Services continuously extract, replicate, and load producer system information in real-time to keep the producer system data with the data in the source system. Since the Data Streaming Services is continuously refreshing, it is able to identify changes within a topic and update its log files with the corrected information. The real-time data streaming feature in middleware tools ensure that any changes within a producer system are automatically reflected in the consumer system when a topic is queried. The source systems' PIAs detail the opportunities that USCIS customers have to correct their PII during the immigration and naturalization process.

2.5 Privacy Impact Analysis: Related to Characterization of the Information

Privacy Risk: There is a risk that inaccurate information is transferred using the Data Streaming Services.

Mitigation: This risk is not mitigated. The Data Streaming Services depends on the accuracy and quality of information from each source system (producer system to consumer system). The Data Streaming Services does not change data "in route" to the receiving system other than to provide standardized formatting of the data, such as date and time formatting. The transactional logs are routinely reviewed to ensure that the data remains accurate and complete during the transmission of data from the producer system to the consumer system.

Privacy Risk: There is a risk of exposing or retaining personal information in audit data.

Mitigation: The Data Streaming Services mitigate this risk by not retaining operational data in the audit log files. Auditing is a fundamental security principle that provides the ability to track the activities of a user or system that may access information maintained within a system. Audit trails track the identity of each subject attempting to access a system, the time and date of access, and the time of log off. Data in the audit log files contain general transactional information that is helpful in identifying the transaction and user.

Privacy Risk: There is a risk that inaccurate information is transferred to the end systems from the Data Streaming Services.

Mitigation: The Data Streaming Services do not change the data "in route" from the producer to the consumer other than to provide standardized formatting of the data, such as date and time formatting.

Privacy Risk: There is a risk that the Data Streaming Services maintains an over collection of data.

Mitigation: This risk is not mitigated. The risk of maintaining an excess of USCIS information is not a risk inherent to Data Streaming Services, which only serves as the conduit to moving data from one place to another. The risk lies in the producer systems that provide data to the consumer systems via Data Streaming Services (such as CMS[11] or myUSCIS,[12] respectively). By design, the Data Streaming Services extract, replicate, and load large amounts of data from USCIS, DHS, and external systems. Data Streaming Services access and send large amounts of duplicated data to efficiently support both USCIS mission-related operations as well as the human resources function of USCIS by providing information for reporting and analytics, statistical analysis, and adjudication purposes. Data Streaming Services would not be considered a valuable

[11] The forthcoming USCIS Content Management Services PIA, will be *available at* www.dhs.gov/privacy.
[12] *See* DHS/USCIS/PIA-071 myUSCIS Account Experience, *available at* www.dhs.gov/privacy.

tool without the ability to access and send a large amount of data between systems. However, it does not alter the data "in route" from the producer to the consumer systems.

Section 3.0 Uses of the Information

3.1 Describe how and why the project uses the information.

USCIS uses Data Streaming Services to transport data between a source system (i.e., producer) and a destination system (i.e., consumer). As a middleware that provides a data transport service, the Data Streaming Services do not have user interfaces. However, USCIS generally uses the Data Streaming Services to support efficient and effective adjudication and administration of immigration benefit requests. Specific use of data by the consumer system is outlined in the respective PIA, and noted in the appendices to this PIA.

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?

Data Streaming Services are intermediary back-end tools that do not have a user interface. Data within the Data Streaming Services is accessible via a system connection only. Data Streaming Services are expected to be used exclusively by USCIS systems and does not have partnership roles with other components.

3.4 Privacy Impact Analysis: Related to the Uses of Information

Privacy Risk: There is a risk that users may inappropriately use the information transported through Kafka and various commercial middleware tools.

Mitigation: Kafka and the middleware tools serve as a repository of data for the purpose of reporting and archiving. There is no direct access to Kafka and middleware tools other than for system administrator purposes. To access Kafka and middleware tool information, users must access the information by accessing an interconnected system. For example, if a user has access to CLAIMS 3,[13] he or she is able to access the data streamed into CLAIMS 3 by Kafka and the

[13] *See* DHS/USCIS/PIA-016(a) CLAIMS 3 and Associated Systems, *available at* https://www.dhs.gov/privacy.

designated middleware tool from the producer system. Furthermore, access controls are in place that determine what data a user of an interconnected system can access. To ensure these access controls, USCIS executes an Interface Control Agreement between Kafka and designated middleware tool and each connected system to define the data the connected system may retrieve from Kafka and the designated middleware tool.

Kafka and the various commercial middleware tools have a limited number of dedicated users. These users include system administrators with privileged system accounts who access Kafka and middleware tool databases for system maintenance purposes. Browsing data within Kafka and the various commercial middleware tools while administering the system is prohibited by USCIS policies.

Privacy Risk: There is a risk that there are no consistent protocols for adding and removing topics, producers, and consumers.

Mitigation: This risk is mitigated through operational procedures established by USCIS Office of Information Technology. USCIS has standard review and approval processes in place for adding and removing new topics, producers, and consumers. Any changes to a system must go through the USCIS change control process and the proper approving authorities before adding and removing new topics, producers, and consumers. USCIS also updates its system documentation to keep an accurate inventory of current topics, producers, and consumers.

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.

USCIS provides general notice to individuals through a Privacy Notice on the instructions to all USCIS forms, which are the original point of collection. This PIA, associated source system PIAs, and SORNs listed in the appendices to this PIA also provide notice.

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

USCIS provides applicants seeking USCIS benefits with a Privacy Notice contained on all benefit request form instructions. The Privacy Notice details the authority for the collection of the information requested and the uses of information. As a general rule, USCIS provides notice that the information collection is voluntary, and that the individual may decline to provide the requested information. However, failure to provide the requested information may delay a final decision or

result in the denial of the applicant's immigration request. On each immigration request form, USCIS includes a release authorization statement that requests the applicant's signature to permit USCIS to release any information from the applicant's records necessary to determine eligibility for the requested benefit.

4.3 Privacy Impact Analysis: Related to Notice

Privacy Risk: Because Kafka and the various commercial middleware tools are not the source system of collection, there is a risk that individuals will not receive notice of the purpose for which Kafka and the various commercial middleware tools uses their information.

Mitigation: USCIS provides the individual with a Privacy Notice explaining the purpose of collection at the original point of collection, and notice of source system interactions with Kafka and the various commercial middleware tools are provided through the publication of this PIA, the applicable source system PIAs, and the applicable SORN(s).

Section 5.0 Data Retention by the project

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

USCIS retains the user and audit logs associated with these services in accordance with General Records Schedule DAA-GRS2013-0006-0003, which states that the records are destroyed when the business use ceases. USCIS identified that it has the need to maintain audit logs for seven days to ensure the successful transport of data from producer to consumer.

5.2 Privacy Impact Analysis: Related to Retention

There is no risk related to retention of data because the Data Streaming Services do not save data, they simply transport data from producers to consumers.

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.

Kafka and the various commercial middleware tools do not share information outside of DHS.

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

Not applicable.

6.3 Does the project place limitations on re-dissemination?

Not applicable.

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

Not applicable.

6.5 Privacy Impact Analysis: Related to Information Sharing

Not applicable.

Section 7.0 Redress

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

An individual may gain access to his or her USCIS records by filing a Privacy Act or Freedom of Information (FOIA) request. Only U.S. citizens, lawful permanent residents, and individuals covered by the Judicial Redress Act of 2015 (JRA) may file a Privacy Act request.[14] Any person, regardless of immigration status, may file a FOIA request. If an individual would like to file a Privacy Act or FOIA request to view his or her USCIS record, he or she may mail the request to the following address:

U.S. Citizenship and Immigration Services
National Records Center
FOIA/PA Office
P O Box 648010
Lee's Summit, MO 64064-8010

Further information about Privacy Act/FOIA requests for USCIS records is available at http://www.uscis.gov.

[14] The Judicial Redress Act of 2015, 5 U.S.C. § 552a note, extends certain rights of judicial redress under the Privacy Act to citizens of certain foreign countries or regional economic organizations; more information is *available at* https://www.justice.gov/opcl/judicial-redress-act-2015.

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

Data Streaming Services do not employ any mechanisms that allow individuals to amend erroneous information. Data Streaming Services maintains read-only data obtained from the source systems, and USCIS personnel cannot amend Data Streaming Services records directly. Data Streaming Services has a refresh mechanism that updates Data Streaming Services on a regular basis to reflect any changes in the source systems records; this refresh helps ensure timely and accurate data.

While Data Streaming Services does not permit individuals to correct inaccurate or erroneous information itself, U.S. citizens, lawful permanent residents, and other persons with records covered by the JRA are afforded the ability to correct information within source systems and interconnected systems by filing a Privacy Act Amendment request under the Privacy Act. U.S. citizens, lawful permanent residents, and persons covered by the JRA should submit requests to contest or amend information contained in Data Streaming Services producer systems and consumer systems as discussed in Section 7.1. The requestor should clearly and concisely state the information being contested, the reason for contesting or amending it, the proposed amendment, and any evidence of the correct information. The record must be identified in the same manner as described for making a request for access. If the request is accepted, any amendment would only apply to USCIS-held information. Persons not covered by the Privacy Act are not able to amend their records through FOIA. If non-U.S. persons find inaccurate information in their records received through FOIA, they may visit a local USCIS Field Office to identify and amend inaccurate records with evidence supporting their reasons for amendment.

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

Data Streaming Services does not employ mechanisms or procedures to notify individuals on how to amend their information that may be contained within Data Streaming Services. Data Streaming Services extract, load, and transport data between USCIS systems. USCIS corrects an individual's information by updating the source system that initially collected the information. This corrected information will then populate Data Streaming Services. The SORNs and PIAs for the source systems explain how individuals can correct erroneous information.

7.4 Privacy Impact Analysis: Related to Redress

Privacy Risk: There is a risk that USCIS may not afford an individual adequate opportunity to correct information retrieved by Data Streaming Services from the connected IT systems.

Mitigation: Data Streaming Services are not the system of record for any of its stored or transferred data. Data Streaming Services provides a mechanism to access and share data to and from multiple connected systems. The information accessed and retrieved by Data Streaming Services is obtained from connected IT systems. The underlying connected IT systems are fully responsible for any information sent to or provided by Data Streaming Services. It is the responsibility of the connected system owner to provide procedures for access and redress in accordance with FOIA/PA. Individuals may seek more information on access, redress, or correction by reviewing the PIA for the individual system.

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?

USCIS ensures information is used in accordance with the stated practices in this PIA by implementing security controls to safeguard PII. Kafka and the various commercial middleware tools are housed in the FedRAMP-approved Amazon Web Services (AWS) U.S. East/West cloud environment, at a moderate confidentiality that allows USCIS to host PII.[15] AWS U.S. East/West is a multi-tenant public cloud designed to meet a wide range of regulatory requirements, to include Government compliance and security requirements.[16] FedRAMP is a U.S. Government-wide program that delivers a standard approach to the security assessment, authorization, and continuous monitoring for cloud services.

USCIS employs technical and security controls to preserve the confidentiality, integrity, and availability of the data, which are validated during the security authorization process. These technical and security controls mitigate the privacy risks associated with unauthorized access and disclosure. The complete list of controls is included in the Kafka and the various commercial middleware tools System Security Plans. Security measures are in place to ensure that the PII in Kafka and the various commercial middleware tools is accessed and used in accordance with DHS and USCIS policies and guidelines.

USCIS only grants back-end access to Kafka and commercial middleware tools to authorized personnel on a strictly need-to-know basis. USCIS audits user access in accordance with the DHS Sensitive Systems Policy Directive, which requires auditing capabilities that log the activity of each user in order to reduce the possibility of misuse and inappropriate dissemination

[15] https://marketplace.fedramp.gov/#/product/aws-us-eastwest?status=Compliant&sort=productName.
[16] Public clouds are owned and operated by third-party service providers whereas private clouds are those that are built exclusively for an individual enterprise.

of information.[17] All user actions are tracked via audit logs to identify information by user identification, network terminal identification, date, time, and data accessed. All USCIS systems employ auditing measures and technical safeguards to prevent the misuse of data.

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

All federal employees and contractors are required to complete annual privacy awareness and computer security awareness training to ensure their understanding of properly handling and securing PII. The Privacy Awareness training addresses appropriate privacy concerns, including Privacy Act obligations (e.g., SORNs). The Computer Security Awareness training examines appropriate technical, physical, personnel, and administrative controls to safeguard information. USCIS also provides role-based training on the proper uses of USCIS information.

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

Only Kafka and the various commercial middleware tools' System Administrators, Database Administrators, and select Subject Matter Experts have direct access to Kafka and the various commercial middleware tools. USCIS audits their access on an operating system level. Users of interconnected systems cannot access Kafka and the various commercial middleware tools directly. Rather, users access Kafka and the various commercial middleware tools information through the systems to which Kafka and the various commercial middleware tools provide data.

Only USCIS users with a valid need-to-know and an account can access Kafka and the various commercial middleware tools log into the system to receive data from Kafka and the various commercial middleware tools. All Kafka and the various commercial middleware tools-connected systems have system logs that indicate successful and failed logins and details about conducted queries. These systems can also audit the list of registered users and each user's access permissions. Auditing does not include the ability to identify specific records each user can access, but audits can include the ability to identify the subject areas that each user accessed. Some systems conduct self-audits by reviewing logs, permissions, and access forms.

[17] *See* DHS Sensitive Systems Policy Directive 4300A (2014).

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?

Kafka itself does not share information outside of USCIS. Kafka, as a data provider to USCIS interconnected systems, may share information routinely or on an ad hoc basis through a Computer Readable Extract (CRE).[18] If USCIS shares information routinely though the source system, an MOU between USCIS and the receiving agency is in place. If sharing occurs on an ad hoc basis through a CRE, USCIS follows the DHS 4300A Sensitive System Handbook - Attachment S1 - Managing CREs containing SPII. Within USCIS, all of Kafka's source systems (listed in the appendices to this PIA) have an Interface Control Agreement with Kafka itemizing each data element that is transmitted between the systems.

8.5 Privacy Impact Analysis: Related to the Accountability and Integrity of the Information

Privacy Risk: The data maintained by Amazon Web Services (AWS) for the purposes of cloud hosting may be vulnerable to breach because security controls may not meet system security levels required by DHS.

Mitigation: This risk is mitigated. USCIS is responsible for all PII associated with the data streaming service, whether on USCIS infrastructure or on a vendor's infrastructure and it therefore imposes strict requirements on vendors for safeguarding PII data. USCIS strictly adheres to the DHS 4300A Sensitive Systems Handbook, which provides implementation criteria for the rigorous requirements mandated by DHS's Information Security Program.[19] USCIS cloud service providers must be FedRAMP-certified. By using FedRAMP-certified providers, USCIS leverages cloud services assessed and granted provisional security authorization through the FedRAMP process to increase efficiency while ensuring security compliance. All contracted cloud service providers must follow DHS privacy and security policy requirements. Before using AWS, USCIS verified through a risk assessment that AWS met all DHS privacy and security policy requirements. Further, all cloud-based systems and service providers are added to the USCIS Federal Information Security Modernization Act (FISMA) inventory and are required to undergo a complete security authorization review to ensure security and privacy compliance. As part of this process, the DHS

[18] DHS defines ad hoc CREs as unplanned, one-time data retrieval events created in response to a specific need for information and not otherwise previously authorized by management or covered in the source system's security plan or by an established Information Sharing and Access Agreement (ISAA).
[19] See https://www.dhs.gov/publication/dhs-4300a-sensitive-systems-handbook.

Senior Agency Official for Privacy reviews all FedRAMP cloud service providers for privacy compliance and privacy controls assessments as part of the privacy compliance review process.

Responsible Official

Donald K. Hawkins
Privacy Officer
U.S. Citizenship and Immigration Services
Department of Homeland Security

Approval Signature

Original signed copy on file with DHS Privacy Office

_____

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

APPENDIX A

All USCIS systems exchange information through various data streaming services. The data streaming services are a combination of data delivery tools and connections to facilitate the seamless communication between different USCIS systems. Producers do not directly connect to Consumers, but rely on the data streaming services to share information. USCIS uses data streaming services to integrate existing systems with new applications and support services. The integration with these data streaming services allows Producers and Consumers to share and receive information from other systems without adversely impacting the availability of each system.

Producer	Consumers
myUSCIS (not an acronym)[20]	• Computer Linked Application Information Management Systems (CLAIMS 3)[21] • Investor File Adjudication Case Tracker (INFACT)[22] • Global (not an acronym)[23] • USCIS Electronic Information System (USCIS ELIS)[24] • Content Management System (CMS) via STACKS (not an acronym)[25] • Enterprise Correspondence Handling Online (ECHO)[26]
CMS-STACKS	• CLAIMS 3 • Global • INFACT • USCIS ELIS

[20] *See* DHS/USCIS/PIA-071 myUSCIS Account Experience, *available at* www.dhs.gov/privacy.
[21] *See* DHS/USCIS/PIA-016(a) CLAIMS 3 and Associated Systems, *available at* www.dhs.gov/privacy.
[22] *See* forthcoming USCIS Investor Immigrant Program PIA, *available at* www.dhs.gov/privacy.
[23] *See* DHS/USCIS/PIA-027(c) USCIS Asylum Division and associated updates, *available at* www.dhs.gov/privacy.
[24] *See* DHS/USCIS/PIA-056 USCIS ELIS and associated updates, *available at* www.dhs.gov/privacy.
[25] The forthcoming USCIS Content Management Services PIA, will be *available at* www.dhs.gov/privacy.
[26] *See* DHS/USCIS/PIA-063 Benefit Decision and Output, *available at* www.dhs.gov/privacy.

	• Freedom of Information Act (FOIA) Immigration Records System (FIRST)[27] • ECHO
CLAIMS 3	• myUSCIS • CMS-STACKS
Global	• myUSCIS • CMS-STACKS
INFACT	• myUSCIS • CMS-STACKS
USCIS ELIS	• myUSCIS • CMS-STACKS

[27] *See* DHS/USCIS/PIA-077 Freedom of Information Act (FOIA) Immigration Records System (FIRST), *available at* www.dhs.gov/privacy.