



Privacy Impact Assessment
for the

Eligibility Risk and Fraud Assessment Testing Environment

April 9, 2010

Contact Point

Glenn Norton

Office of Transformation Coordination

Reviewing Official

Donald Hawkins

Chief Privacy Officer

U.S. Citizenship and Immigration Services

(202) 272-8000

Mary Ellen Callahan

Chief Privacy Officer

Department of Homeland Security

(703) 235-0780



Abstract

The Office of Transformation Coordination of the United States Citizenship and Immigration Service is planning to prototype the Eligibility Risk and Fraud Assessment Testing Environment. This environment will be used to develop, test, and refine the risk and fraud business rules against historical data extracts before deploying to a full production environment. This new testing involves the use of personally indefinable information.

Overview

The Department of Homeland Security (DHS) through the United States Citizenship and Immigration Service (USCIS) implements immigration law and policy through the processing and adjudication of applications and petitions submitted for citizenship, asylum, and other immigration benefits. These immigration benefits are highly sought after not only by those who legitimately qualify for the benefit, but also by those who do not qualify. USCIS supports the Department's national security mission by preventing individuals from fraudulently obtaining immigration benefits and by denying applications from individuals who pose a national security or public safety threat to the U.S.

Office of Transformation Coordination

USCIS established the Office of Transformation Coordination (OTC) to embark on an enterprise-wide "Transformation Program" that will transition the agency from a fragmented, form-centric, and paper-based operational environment to a centralized, person-centric, consolidated environment utilizing electronic adjudication. The ability to accurately establish risk and fraud trends is a key tool to support this transformation effort. To this end, the new systems that make up this transformed environment that USCIS will deploy will include Eligibility Risk and Fraud Management tools to protect national security. This Privacy Impact Assessment (PIA) covers the prototype test system that will precede the deployment of the Eligibility Risk and Fraud Management tools and will allow for a detailed and more granular description of USCIS emerging risk and fraud technologies and their potential privacy impacts prior to the rule sets being deployed into a production environment.

USCIS is developing and deploying this transformed environment over the next several years through a series of "releases" that deploy discrete information technology capabilities in each release. Each release builds on the previous capabilities to expedite subsequent deployments and grow the maturity of the transformed environment incrementally. Each release will include risk and fraud management capabilities appropriate to the functionality being deployed in the release. The first release will primarily deploy capabilities to support person-centric accounts. These accounts will allow USCIS to facilitate more customer-friendly transactions, such as providing web-based services to allow the account holder to maintain core biographical information, change an address, and reschedule appointments. Person-centric accounts will allow the customer to establish and maintain a unique account with USCIS, providing the agency with the ability to manage customer interactions holistically rather than on a case-by-case basis. Subsequent releases will focus on USCIS lines of business (i.e., nonimmigrant, immigrant,



humanitarian, and citizenship) by deploying additional case management functionality approximately every six months.

Eligibility Risk and Fraud Management Tools

The prototype will allow USCIS to find the best way to systematize and normalize rules consistently across all applicants, reducing the opportunity for human error or misapplication of rules in ways that sometimes differ from one adjudicator to the next. The Eligibility Risk and Fraud Management tools will translate the knowledge of risk and fraud indicators identified by USCIS adjudicative staff into business rules that can be coded into a system. This will be accomplished by implementing business rules within the system to identify areas of concern for USCIS staff who will then be able to review and evaluate those areas of concern as part of the adjudication of immigration benefits. The risk rules will help to identify whether an individual meets the eligibility requirements for a particular benefit. Eligibility requirements include such activities as checking to determine whether the person has been in or out of the country, poses a national security threat, has committed a crime that disqualifies him (as revealed by the existing USCIS background check process), or if historical immigration data matches current application submissions. In addition to the eligibility risk rules, the system will run related rules to identify possible fraud, national security concerns, as well as possible inaccurate information. The combination of the eligibility risk and fraud tools will provide a comprehensive fraud and national security assessment to the adjudicator, who determines whether to approve or deny the benefit application. In order to ensure that these business rules reliably identify eligibility risk and fraud concerns, USCIS has established a “Eligibility Risk and Fraud Assessment Testing Environment” to develop, test and refine these rules before deploying them as part of the transformed environment. This prototype is the first step in the creation of this environment.

The prototype project will help USCIS staff understand and address the risks and benefits of implementing Eligibility Risk and Fraud Management tools in the transformed environment. The prototype’s primary goal is to assess whether the Eligibility Risk and Fraud Management tools can effectively and fairly assess the eligibility risk and fraud issues associated with immigration benefit applications before deploying into a production environment.

The prototype will use two types of analysis: link analysis and rules-based analysis. The link analysis software will attempt to establish certain relationships between individuals in the data extracts. Link analysis can be used to: 1) identify potential duplicate accounts and inconsistent data entries for individuals; 2) establish and maintain links between individuals named on applications or petitions, which can help to assess the risk associated with family-based applications and broad-based immigration benefit fraud conspiracies; 3) map non-obvious relationships using common data on two or more records to link individuals that would not have been possible otherwise; and 4) to produce a visual representation of relationships for any given individual. While these “relationships” will only be used to discover known eligibility risk and fraud patterns, it is important to note that these relationships will need validation by government staff. The result of running link analysis on completed-case data by using the link analysis software is a list of disclosed and undisclosed/non-obvious relationships and connections.

The rules-based analysis will take information identified by USCIS adjudicators and fraud detection and national security officers as indications of either ineligibility, fraud, or national security



threat. The scenario-based rules run against the link analysis and the records to identify possible matches. Scenario-based rules will be developed on both basic eligibility requirements, such as length of stay in the United States, as well as national security and fraud-related activities.

As part of the prototype, any results from running the rules will not be actionable (e.g., as all cases will have been previously adjudicated and closed, any results from the prototype will not affect case disposition). However, if the prototype uncovers instances of widespread fraud, that can be validated, in the immigration process or evidence of a large-scale national security threat, the OTC will forward the results to the Fraud Detection and National Security Directorate (NSRV) for further review. Once this functionality is deployed into a production environment, it will not be used to determine benefit eligibility, but will identify areas of concern for USCIS Adjudicative Staff to review and evaluate before making benefit decisions. This will be accomplished by translating the knowledge of risk and fraud indicators identified by USCIS adjudicative staff into business rules that can be coded into a system that will benefit all USCIS adjudicative staff. It is expected that as USCIS begins to block access to immigration benefits by those that do not qualify for those benefits, those that fraudulently apply, or assist those who do, will begin to adapt and change the methods that they use. To address this concern, the Eligibility Risk and Fraud Testing Assessment Environment will be a long term project that will exist to test the results of new business rules that are developed to address new and evolving risk and fraud trends before deploying to a production environment.

Data Sources

The prototype will use a combination of real USCIS completed-case data, synthetic data from USCIS to simulate misspellings and name variants, and synthetic data from Customs and Border Protection (CBP) that is not about actual individuals. USCIS will use the prototype to develop and assess the results of a risk and fraud analysis tool against a data set of USCIS case information that remains consistent over time, providing the agency the ability to validate test results from a known data set.

Before deploying each release, USCIS plans to develop, test and refine the risk and fraud business rules before those rules are deployed as part of a release. The first step in testing these business rules is to establish a prototype of the Eligibility Risk and Fraud Assessment Testing Environment. A prototype refers to the original instance of something that affords designers the opportunity to test the design, underlying theories, and performance prior to developing a new product. This prototype will be executed in two phases. First, the agency will implement the prototype using a one-time export of completed-case data (e.g., cases that have already been adjudicated and are now closed) contained in the USCIS Fraud Detection and National Security Data System (FDNS-DS). In the second phase of the prototype, another one-time export of completed-case data will be included from the Computer Linked Adjudication Information System (CLAIMS) 3, CLAIMS 4, the Refugee Asylum and Parole System (RAPS) and the Central Index System (CIS), which are all USCIS systems.

The data extracts from FDNS-DS, CLAIMS 3, CLAIMS 4, RAPS, and CIS will consist of completed-case information that has already been adjudicated and closed. This information will be loaded into the prototype system at a USCIS data center and will be used to test, validate, and refine risk and fraud business rules to support the development of the new transformed environment. Access to this system will be restricted to a limited number of USCIS government and contract staff from the OTC who



are working to develop the business rules. These business rules translate the knowledge of eligibility risk and fraud indicators identified by USCIS adjudicative staff into business rules that can be coded into a system. The agency will compare the results obtained from executing these business rules against historical USCIS information to evaluate the effectiveness of the rules in ranking the potential risk and fraud of a particular individual. As part of developing a business rule, OTC will develop a process to ensure that relevant parties, including USCIS and the DHS Privacy Office are consulted prior to deploying a business rule to ensure that pertinent privacy considerations are measured during this evaluation process.

The prototype will not collect any new information, including PII data, and will not disseminate any information that it uses and maintains. The agency will acquire the information for the prototype through a one-time export of completed-case data contained in the CLAIMS 3, CLAIMS 4, RAPS, FDNS-DS and CIS systems. The CLAIMS 3 and CLAIMS 4 PIA's were published on September 5, 2008 by the DHS Office of Privacy, the RAPS PIA on November 24, 2009, the FDNS-DS PIA on July 29, 2008, and the CIS PIA on June 22, 2007. The SORN's for these systems can be found in the Federal Register. (CLAIMS 3 and CLAIMS 4 73 FR 56596, RAPS 75 FR 409, FDNS-DS 73 FR 48231 and CIS 72 FR 1755.) Also, OTC will create synthetic data and it will also be loaded into the system to test specific rules, such as adding a new record with a similar name and the same date of birth to an existing record to see if a link between the two is identified when running the rule.

In addition, USCIS has coordinated with the CBP agency for an export of synthetic data from the TECS system. This data resembles the types of data housed in TECS but does not include any real completed-case data or real personally identifiable information (PII). USCIS is interested in two types of data maintained in the TECS system: 1) arrival/departure information covered by the DHS/CBP-016 Non-Immigrant Information System and DHS/CBP-007 Border Crossing Information systems of records notice, and 2) subject-based lookouts for persons of interest covered by DHS/CBP-011 TECS Enforcement Records. The arrival/departure data is biographical and travel data collected from persons entering and exiting the U.S. used to determine eligibility for certain benefits. TECS enforcement records include lookout information that helps USCIS identify whether an individual may pose a national security risk or are otherwise not eligible for a particular benefit.

The data from all these systems will be encrypted and exported onto a government-issued portable media device at a DHS data center and then loaded into the prototype by a USCIS employee at the same data center. The information will never physically leave the DHS data center and the data on the portable media will be destroyed immediately once the prototype is loaded. This export process complies with the DHS Sensitive Systems Security Policy 4300A. The prototype will not share data with any other system, nor have any connections external to it other than to load the data export.

The prototype will be used for a six month period, in advance of establishing a more comprehensive testing environment. An updated or new PIA will be completed to address this testing environment that will exist long-term so the agency can test rules before deploying to a production environment. Additionally, a PIA will be conducted on the use of the eligibility risk and fraud assessment tool on active applicants.



Section 1.0 Characterization of the Information

1.1 What information is collected, used, disseminated, or maintained in the system?

The prototype will not collect any new information from the public, but will use the following completed-case data from USCIS data sets: Name, Date of Birth, Country of Birth, Country of Citizenship, Gender, Social Security Number, Alien Number, Marital Status, Family Relationships, Current and Past Address Information, and Current and Past Telephone Information.

The prototype will use synthetic data from different datasets in TECS that is not real data on real individuals. This training data pertains to arrivals into and departures from the U.S., and will include: name, date of birth, country of birth, gender, date of admission, travel document information, and travel information, including air carrier and flights. The training data pertaining to enforcement records will include the above information as well information pertaining to known or suspected violators of law, wanted persons, and persons of interest for law enforcement or counterterrorism purposes. Also, OTC will create synthetic data that will be input into the system to test specific rules, such as adding a new record with a similar name and the same date of birth of an existing record to see if a link between the two is identified when running the rule.

The system will also have the eligibility- and fraud-based rules, any new links identified by the system and validated by a USCIS employee, and any flags indicating possible ineligibility, fraud, or national security threats that the system identified.

1.2 What are the sources of the information in the system?

The prototype will use and maintain completed-case data from CLAIMS 3, CLAIMS 4, RAPS, FDNS-DS, and CIS, along with synthetic data from DHS/CBP-016 Non-Immigrant Information System, DHS/CBP-007 Border Crossing Information, and DHS/CBP-011 TECS Enforcement Records. Also, OTC will create synthetic data that will be input into the system to test specific rules, such as adding a new record with a similar name and the same date of birth to an existing record to see if a link between the two is identified when running the rule.

1.3 Why is the information being collected, used, disseminated, or maintained?

The information in the prototype is being used and maintained in order to test, validate, and refine risk and fraud business rules to support the development of a new environment that consistently applies these eligibility risk and fraud related rules to all applications for immigration benefits. USCIS will use this information to develop and assess the results of a risk and fraud analysis tool against a static set of historical data before the business rules are deployed into any operational environment. This will allow for an assessment of privacy impacts prior to the rule sets being deployed into a production environment. What USCIS learns through the prototype process will then better inform the longer term solution and will be addressed in future PIAs.



1.4 How is the information collected?

The agency will acquire the information for the prototype through a one-time export of data contained in the CLAIMS 3, CLAIMS 4, RAPS, FDNS-DS and CIS systems of records. The data will be encrypted and exported into a government issued portable media device at a DHS data center and then will be loaded into the prototype by a USCIS employee at the same data center. The information will never physically leave the DHS data center and the data on the portable media will be destroyed immediately once the prototype is loaded. This export process complies with the DHS Sensitive Systems Policy 4300A. The information maintained in this prototype is collected through methods established in the corresponding Systems of Record Notices.

In addition, USCIS has coordinated with the CBP agency for an export of test data from the TECS system. This test data does not including any real data.

Also, OTC will create synthetic data will be input into the system to test specific rules, such as adding a new record with a similar name and the same date of birth to an existing record to see if a link between the two is identified when running the rule.

1.5 How will the information be checked for accuracy?

The information loaded into the system for testing the prototype will not have additional checks for accuracy, but will rely on the accuracy validation process of the system from which it originates. The prototype is designed to test, validate, and refine risk and fraud business rules, on data similar to what the agency will receive as part of the application process and updated as part of the benefit screening process, in order to create a better risk and fraud detection mechanism. Part of this process will include identifying information that is possibly inaccurate.

As USCIS utilizes the prototype it will establish relationships between the various data that is loaded into the system. These new relationships will be used to validate the information provided by the individual and identify whether there are possible eligibility risks and fraud.

1.6 What specific legal authorities, arrangements, and/or agreements defined the collection of information?

The authority for maintenance of records in the system is found in 8 U.S.C. 1324a, 8 U.S.C. 1360, 42 U.S.C. 1320b-7 and the Immigration Reform and Control Act of 1986 (IRCA), Public Law (Pub. L.) 99-603, The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA), Public Law 104-193, 110 Stat. 2168, Title IV, Subtitle A, of the Illegal Immigration Reform and Immigrant Responsibility Act of 1996 (IIRIRA), Public Law 104-208, 110 Stat. 3009, 18 U.S.C. 3291, and in Executive Order 12989, as amended by Executive Order 13465, June 6, 2008.



1.7 Privacy Impact Analysis: Given the amount and type of data collected, discuss the privacy risks identified and how they were mitigated.

Any risks associated with the data to be used in the prototype are significantly mitigated by the fact that the prototype is a test environment with a primary goal of ferreting out the operational and privacy impacts associated with the data before deploying into production mode. There are two primary risks associated with the data to be used in the prototype. First, is the risk that the data extracts will be lost, duplicated or not destroyed properly. To mitigate this risk, USCIS is following standard procedures documented as part of our Systems Engineering Life Cycle (SELC) methodology and these procedures are compliant with DHS Sensitive Systems Policy 4300A. The second identified risk is that the data once it has been loaded will be used for other purposes not covered as part of this PIA. To mitigate this risk the prototype will be conducted in a closed environment. Also, the information included through data extracts will be introduced to the prototype environment in stages, as the rules development and test schedule dictate, which will minimize the amount of information and its exposure. In addition the results of running the business rules in the prototype will not be actionable. The testing of closed-case data will help refine risk and fraud indicators into business rules; as the case has already been adjudicated and closed, the data will not be used in adjudicative decisions, but a particularly egregious fraud or threat pattern may be forwarded to NSRV to mitigate against privacy risks. Since the TECS synthetic data is not valid, no national security risks could be identified on any subject based query information. It is intended that the information acquired, be used and maintained in the prototype be the minimal amount of information needed to test, validate, and refine risk and fraud business rules before they are implemented. Given the wide range of types of data used and maintained in the prototype, the agency will take measures to mitigate privacy concerns, including limiting access to the prototype to those with a need to know.

USCIS has determined that the privacy risks associated with using such a large extract of personally identifiable case data are justified by the fact that a large volume of data is required for this working prototype environment. This is necessary since the underlying software that comprises the eligibility risk and fraud tool works differently based on the amount of data loaded into the system. The large amount of data loaded closely replicates the amount of data the final transformation Integrated Operating Environment will contain and will provide the opportunity to assess how the system will operate. Likewise, USCIS is purposefully not making an effort to update or perfect the data before testing, because the intention is to test real data that may contain inaccuracies in order to glean useful test results. For this reason, USCIS would not act on any of the results of the tests, except in possible egregious circumstances where data may be referred over to trained fraud or national security personnel for their individual analysis and input.



Section 2.0 Uses of the Information

2.1 Describe all the uses of information.

The program will use the combination of historical closed case data, TECS training data and synthetic data to test, validate, and refine risk and fraud business rules before they are implemented into the production environment Risk Analyzer. The prototype will be used to test individual entity relationship mapping, a process by which the prototype will identify potential duplicate individuals within the USCIS environment. For example, if there is a record that contains a variety of information on a William Smith, another for a Bill Smith, and yet another for a Billy Smith, the prototype will be tested to see if the system can identify if all these records are potentially the same person. This will be done by looking at the other information that USCIS has on the individuals like date of birth, country of birth etc.

2.2 What types of tools are used to analyze data and what type of data may be produced?

The prototype will use link and rule-based analysis to assess the risks and identify the reliability of previously known and unknown fraud indicators.

Link analysis will establish relationships between individuals in the data extracts. It can be used to identify potential duplicate accounts and inconsistent data entries for individuals; establish and maintain links between individuals named on applications or petitions, which can help to assess the risk associated with family-based applications and broad-based immigration benefit fraud conspiracies; map non-obvious relationships using common data on two or more records to link individuals that would not have been possible otherwise; and to produce a visual representation of relationships for any given individual. While these “relationships” will only be used to discover known risk and fraud patterns, it is important to note that these relationships will need validation by government staff. For example, if there is a record that contains a variety of information on a William Smith, another for a Bill Smith, and yet another for a Billy Smith, the prototype will be tested to see if the system can identify if all these records are potentially the same person. This will be done by looking at the other information that USCIS has on the individuals like date of birth, country of birth etc. This information is then presented to government staff to make a final determination of whether the records are about the same person.

The result of running link analysis on case data by using link analysis software is a list of disclosed and undisclosed/non-obvious relationships and connections. The rule-based analysis tools uses these relationships to create a notification or “flag” of any elements in a subject-based query that reveal connections that a particular risk and fraud rule is designed to reveal. After running the “rule” the applicant’s information would reveal such a link and the software will flag that applicant for additional review by a USCIS adjudicator.



2.3 If the system uses commercial or publicly available data please explain why and how it is used.

The prototype will not use any commercial or publicly available data. Rather, it will only use closed case data obtained from the USCIS systems of records (CLAIMS 3, CLAIMS 4, RAPS, FDNS-DS, and CIS systems), along with TECS training data (i.e. synthetic data.)

2.4 Privacy Impact Analysis: Describe any types of controls that may be in place to ensure that information is handled in accordance with the above described uses.

There are several risks associated with the use of this data. First, there is a risk that the information resulting from the prototype could have an impact on an adjudicative decision, which would be out of compliance with the scope of the prototype. The second risk is that the rules that are developed provide too many false positive hits or provide erroneous links.

To address the first risk of an impact on an adjudicative decision, no results of the testing of the prototype will be actionable. Since the data being used is closed-case data, the case has already been adjudicated and closed and any results from the prototype will not affect case disposition. For the second risk regarding false positive hits or erroneous links, the purpose of the prototype is to fine-tune risk and fraud rule sets so that once they are deployed in a production environment against real data and real cases, the privacy of applicants will not be adversely affected. The prototype environment will operate in a USCIS Data Center with the associated security controls consistent with DHS System Security Handbook 4300A. The prototype environment will be isolated from other systems/environments and provide for controlled and limited access. Extracted data will be loaded in prototype environment and system-to-system connections will not be permitted. Data disposition will be documented in the request for the extract of the data and will include provisions to destroy all data received from the original data extract at the end of the six-month prototype. These controls will be implemented and monitored to provide reasonable assurance and protection of the extracted data to be used only as planned for in the prototype environment. The primary reasons for the implementation of the prototype is to help USCIS staff understand and address the issues associated with the implementation of a Risk Analyzer tool in the transformed environment and understand and explain the role of the tool in the new integrated operating environment. The prototypes will improve USCIS' ability to assess the risk and fraud issues associated with immigration benefit applications before deploying into a production environment, while taking into consideration privacy concerns.

Section 3.0 Retention

3.1 What information is retained?

The prototype will utilize completed case data from CLAIMS 3, CLAIMS 4, RAPS, FDNS-DS, and CIS, along with training data from the CBP TECS system and the synthetic data created by USCIS.



However, no case data will be retained. All data extracts will be destroyed at the end of the six-month prototype period.

3.2 How long is information retained?

The data extracts will be retained while the rules development process requires use of the detailed information, but no longer than the duration of the six-month prototype. The information extracted and supplied to the prototype will be destroyed once testing is complete or within 90 days, whichever is shorter.

3.3 Has the retention schedule been approved by the component records officer and the National Archives and Records Administration (NARA)?

In working with the USCIS Records Officer it was determined that this prototype contains read only data from existing systems and as such is a non-records system and is not required to have a retention schedule set with NARA. Although the agency will not establish a retention schedule with NARA for the prototype system the original systems of record have retention schedules, for the information they retain, that have been published and approved by NARA.

3.4 Privacy Impact Analysis: Please discuss the risks associated with the length of time data is retained and how those risks are mitigated.

From an information retention perspective, the primary risk of this prototype is that retaining the data from longer than necessary increases the risk of unauthorized access, use, and loss of the data. USCIS mitigates this risk by destroying data as soon as it's no longer required as outlined above. This will additionally ensure that USCIS will not inadvertently use or retain information that has exceeded its underlying retention schedule from source systems.

To ensure that data is actually destroyed per the retention plan, USCIS has documented the data disposition in the request for the extracted data and will include provisions to destroy all data received from the original data extracts at the end of the six month prototype. In addition, this is a prototype system that will be used for testing purposes and will not be accessed by other than a limited number of USCIS government and contract staff from the OTC who are working to develop the business rules. These controls will be implemented and monitored to provide reasonable assurance and protection of the extracted data to be used only as planned for in the prototype. These controls are documented as part of the disposition phase in the SELC documentation maintained by the USCIS Office of Information Technology (OIT).



Section 4.0 Internal Sharing and Disclosure

4.1 With which internal organization(s) is the information shared, what information is shared and for what purpose?

The information used and maintained by the OTC as part of the prototype will not be accessible or shared with any office within USCIS, or other DHS component agencies, offices or directorates. The information will be used and maintained solely by the OTC to refine, test, and validate its risk and fraud business rules. *Ad hoc* hits and other results that may identify risk and fraud will not be actionable. However, if the prototype uncovers instances of widespread fraud, that can be validated, in the immigration process or evidence of a large-scale threat pattern, the OTC will forward the results to the NSRV for further review.

4.2 How is the information transmitted or disclosed?

No system to system connectivity will occur to transmit or disclose data.

4.3 Privacy Impact Analysis: Considering the extent of internal information sharing, discuss the privacy risks associated with the sharing and how they were mitigated.

The primary risk associated with the application of business rules run against closed case data is that the results of the prototype may put into question the decision on a case. To mitigate this risk no information will be shared from the system within the agency.

Section 5.0 External Sharing and Disclosure

5.1 With which external organization(s) is the information shared, what information is shared, and for what purpose?

The prototype will not establish any interfaces with any external organization. The information used and maintained by the OTC as part of the prototype will not be shared with any external office to DHS. The information will be used and maintained solely by the OTC to refine, test, and validate its risk and fraud business rules. If the prototype uncovers instances of widespread fraud, that can be validated, in the immigration process or evidence of a large-scale threat pattern threat, the OTC will forward the results to the Fraud Detection and NSRV for further review. If NSRV determines that there is a threat that needs further investigation or action they may choose to share the information with the appropriate external organization.



5.2 Is the sharing of personally identifiable information outside the Department compatible with the original collection? If so, is it covered by an appropriate routine use in a SORN? If so, please describe. If not, please describe under what legal mechanism the program or system is allowed to share the personally identifiable information outside of DHS.

This section is not applicable as no information will be shared with an external organization. In the event that NSRV determines that information needs to be shared from the system because of an egregious instance of fraud or threat pattern, that information would be shared in compliance with the routine use given in the System of Records Notice (SORN) that applied in that instance.

5.3 How is the information shared outside the Department and what security measures safeguard its transmission?

This section is not applicable as no information will be shared with an external organization.

5.4 Privacy Impact Analysis: Given the external sharing, explain the privacy risks identified and describe how they were mitigated.

This section is not applicable as no information will be shared with an external organization.

Section 6.0 Notice

6.1 Was notice provided to the individual prior to collection of information?

For this prototype, the information is not collected directly from individuals and is used only for internal testing, so USCIS relies on the notice for the underlying initial collection. Individuals whose information will be used in the prototype were provided notice as part of the process of applying for a benefit with USCIS or legacy INS. Because the prototype only uses completed case data and does not collect any new information, it does not provide notice beyond that which the individuals received at the time they applied for one or more benefits. Individuals who apply for USCIS benefits are presented with a Privacy Act Statement as required by Section (e)(3) of the Privacy Act and sign a release authorization on the benefit application/petition. The Privacy Act Statement details the authority to collect the information requested and uses to which USCIS will put information the applicant provides on immigration forms and in support of an application.

In addition to the publication of this PIA, SORN were published for the applicable systems that are providing the data extracts have published SORNs:



- FDNS-DS, which is used to record, track, and manage the background check and adjudicative processes related to immigration applications and petitions with suspected or confirmed fraud, criminal activity, egregious public safety, and / or national security concerns, and cases randomly selected for benefit fraud assessments; 73 FR 48231
- CLAIMS 3 and CLAIMS 4, which track Immigrant, Non-Immigrant, and Naturalization case work and are referenced in the Benefits Information Systems; 73 FR 56596
- RAPS, which is used to process Asylum and Refugee applications; 75 FR 409
- And CIS which contains personal identification data such as A-File number, date and place of birth, date and port of entry as well as the hardcopy paper A-file's physical location. 72 FR 1755.

6.2 Do individuals have the opportunity and/or right to decline to provide information?

No. Because the prototype will only use closed-case data for an internal testing purpose that will not impact the individual's application benefit, there is no additional opportunity to opt out.

6.3 Do individuals have the right to consent to particular uses of the information? If so, how does the individual exercise the right?

The consent to use and maintain this information was given at the time that the applicant submitted his or her benefit application. Because the prototype is not collecting any new data and is simply making use of completed case data for which USCIS has already received consent, the prototype will not offer individuals an additional opportunity to consent to the use of their information. With this secondary use of the information supplied by the individuals, USCIS has taken care to ensure that the results of this prototype will not impact the person who supplied the information. The results generated by the prototype will not be shared with any office within USCIS, or other DHS component agencies, offices or directorates. *Ad hoc* hits and other results that may identify risk and fraud will not be actionable.

As part of the process to apply for immigration benefits the benefit applications require that applicants provide certain information requested in an application. This information is critical in making an informed adjudication decision to grant or deny a USCIS benefit. The failure to submit such information prohibits USCIS from processing and properly adjudicating the application/petition and thus precludes the applicant from receiving the benefit. Therefore, during the application process, individuals consent to the use of the information submitted for adjudication purposes. Specifically, all USCIS immigration forms include a Privacy Act Statement and require the applicant's signature authorizing "the release of any information from my records that USCIS needs to determine eligibility for the benefit." This information is also conveyed in the SORNs for this system and in the Privacy Act Statement on the application itself. Applicants grant consent to the collection and use of the information when they sign the application.



6.4 Privacy Impact Analysis: Describe how notice is provided to individuals, and how the risks associated with individuals being unaware of the collection are mitigated.

Notice to the public is primarily given through PIA and SORNs associated with the systems of origin where the information was collected. For FDNS-DS, that information is collected through the benefit application process and originally stored in the CLAIMS 3, CLAIMS 4, and the RAPS systems. FDNS-DS and CIS have also published SORNs. By signing the application the benefit seeker is aware of and consents to the collection of information. The privacy risk associated with this particular collection of information is that the individual may not be fully aware that their information will be used to conduct an inquiry into benefits eligibility. To further mitigate this risk the prototype will not share data with any other system, nor have any connections external to it other than to load the data export.

Section 7.0 Access, Redress and Correction

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

Individuals may gain access to their USCIS records by filing a Freedom of Information Act (FOIA) or Privacy Act request.

USCIS may elect to withhold any related law enforcement sensitive information relating to a requestor, which could possibly compromise ongoing criminal investigations if released to the requestor, pursuant to the Privacy Act; 5 U.S.C §552a(k)(2).

An individual may file a FOIA or Privacy Act request to view their USCIS record by submitting a written request to the following address:

National Records Center, FOIA/PA Office
P.O.Box 648010
Lees Summit, MO 64064-8010

Further information for FOIA request for USCIS records can also be found at <http://www.uscis.gov>

7.2 What are the procedures for correcting inaccurate or erroneous information?

If an individual would like to correct known erroneous information in their USCIS record, the individual can file a USCIS form directed at changing the specific erroneous information. For example, an applicant/petitioner can change their address by filing a Change of Address form (AR-11). If an applicant/petitioner believes their file is incorrect but does not know which information is erroneous, the applicant/petitioner may file a Privacy Act request as detailed in Section 7.1.



7.3 How are individuals notified of the procedures for correcting their information?

Individuals are notified of the procedures for correcting their information on USCIS forms, the USCIS website and by USCIS personnel who interact with benefit applicants/petitioners.

7.4 If no formal redress is provided, what alternatives are available to the individual?

Normal USCIS procedure for redress is provided to applicants/petitioners as outline in Sections 7.1 and 7.2.

7.5 Privacy Impact Analysis: Please discuss the privacy risks associated with the redress available to individuals and how those risks are mitigated.

The PII data contained within the prototype is obtained from other USCIS case management systems (see Section 1.2), and as such, individuals must address all information access rights to these systems. Previously established procedures for changing biographical information may be followed to correct known erroneous information, for example file an AR-11 form to change an applicant/petitioner's address. If the applicant/petitioner suspects erroneous information but does not know which part of the information is incorrect, the applicant/petitioner can file a FOIA request as detailed in Section 7.1.

USCIS may elect to withhold any related law enforcement sensitive information relating to a requestor, which could possibly compromise ongoing criminal investigations if released to the requestor, pursuant to the Privacy Act; 5. U.S.C. §552a(k)(2).

All privacy risks associated with redress mechanisms are significantly mitigated by the fact that the testing environment is merely a prototype and by design will assess privacy risks prior to the deployment of a live production environment.

Section 8.0 Technical Access and Security

8.1 What procedures are in place to determine which users may access the system and are they documented?

In compliance with federal law and regulations, users who access the prototype will be granted on a need to know basis. This need to know is determined by the individual's current job functions. Users may have read-only access to the information if they have a legitimate need to know as validated by their supervisor and the system owner and have successfully completed all personnel security training requirements. System administrators may have access if they are cleared and have legitimate job functions that would require them to view the information.



Any user needing access must complete a request for access. This application states the justification for the level of access being requested. The requestor's supervisor, the system owner, and the USCIS Office of the Chief Information Officer (OCIO) review this request; if approved, the requestor's clearance level is independently confirmed and the user account is established.

Criteria, procedures, controls, and responsibilities regarding this system will be developed in standard operating procedures (SOP) defining policies and procedures for determining which users may access the system. Additionally, there are several department and government-wide regulations and directives, which provide additional guidance and direction.

8.2 Will Department contractors have access to the system?

Contractors will have access to the prototype under the direction of the USCIS OIT and the OTC. Access is provided to contractors only as needed to perform their duties as required in the agreement between the agency and the contractor and as limited by relevant SOPs. In addition, agency employees and contractors who have completed the system access application process and been granted appropriate access levels by a supervisor are assigned a login ID and password to access the system. These users must undergo federally approved clearance investigations and sign appropriate documentation to obtain the appropriate access levels. Contractors are also required to sign non-disclosure agreements and Rules of Behavior agreements for system access.

8.3 Describe what privacy training is provided to users either generally or specifically relevant to the program or system?

All agency personnel and contractors take the mandatory on-boarding and annual DHS Information Technology security and privacy awareness training. The agency will also require that those with access complete role-based training to emphasize the specific concerns associated with the use of the prototype.

8.4 Has Certification & Accreditation been completed for the system or systems supporting the program?

Certification and accreditation has been accomplished on the DHS owned and operated data center where the prototype will be physically hosted, and it is in compliance with DHS Sensitive Systems Policy Handbook 4300A. In consultation with the Chief Information Security Officer, within the Office of Information Technology and the USCIS Chief Privacy Officer it was determined that there will be no formal Certification and Accreditation package specific to the prototype environment as the prototype is covered as part of the data center's Authority to Operate; however the following controls will be in place to protect the prototype and associated data:

- Security controls will be in place as part of the General Support System supporting the USCIS Local Area Network;



- Additional security and privacy controls will be implemented and monitored to provide reasonable assurance for the protection of the extracted data and prototype environment. These controls will include, but are not limited to, the following:
 - The prototype environment will be isolated from other systems/environments in the CIS data center;
 - Access to the prototype system will be limited and controlled;
 - Extracted data will be loaded in the prototype environment and system to system connections will not be permitted;
 - The plan for data disposition will be documented in extraction request form and will include provisions to destroy all data received from the original data extracts.

8.5 What auditing measures and technical safeguards are in place to prevent misuse of data?

Auditing and historical use tracking will be designed and built into the prototype. The users will be identifiable by their logon information and will be made aware that they are being monitored through logon warning banners. Audit logs will be reviewed on a regular basis and when questionable activities are identified. Auditing will ensure that users will be able to be held accountable for their handling of PII.

Technical and operational security safeguards to control information used and maintained by the prototype include: access controls, use logging, input controls, media labeling, and policies and training for security policies and procedures.

8.6 Privacy Impact Analysis: Given the sensitivity and scope of the information collected, as well as any information sharing conducted on the system, what privacy risks were identified and how do the security controls mitigate them?

The information used and maintained by the prototype will be considered sensitive but unclassified data. Since the prototype will contain significant amounts of PII, this information will be protected by a full array of technical, physical, and operational controls as required by DHS Sensitive Systems Policy Handbook 4300A. These security controls include transmission and at-rest encryption, access controls, auditing, retention limits, logging of users' activities, as well as logging and management of electronic extracts per Office of Management and Budget Memorandum 06-16. Users will be held personally responsible for handling this information appropriately and will be provided procedures and training appropriate for their assigned role and responsibilities.



Section 9.0 Technology

9.1 What type of project is the program or system?

The prototype is a development system that will be used to implement the Eligibility Risk and Fraud Assessment tools as part of the agency's development of a new integrated operational environment. This system will be used to develop, test, and refine the risk and fraud business rules against data extracts imported from USCIS' CLAIMS 3, CLAIMS 4, RAPS, FDNS-DS, and CIS system. Once the risk and fraud business rules have been tested and proven they will be deployed to a full production environment.

9.2 What stage of development is the system in and what project development lifecycle was used?

The prototype as part of establishing the new integrated operation environment is in the development stage of the DHS SELC. This prototype is a six month project activity to establish system configuration and to test the commercial off the shelf (COTS) software in its implementation of the risk and fraud assessment tools work.

9.3 Does the project employ technology which may raise privacy concerns? If so please discuss their implementation.

Yes. The use of link analysis and development of business rules on application data may rise privacy concerns. To address these concerns, USCIS is starting by testing these technologies in a prototype, non-operational environment so that risks to PII are mitigated.

The prototype will make use of three COTS software products to perform risk and fraud analysis. These COTS products will facilitate "link analysis" functionality. The prototype will establish relationships between individuals in the data extracts. The prototype can be used to: 1) identify potential duplicate accounts and inconsistent data entries for individuals; 2) establish and maintain links between individuals named on applications or petitions, which can help to assess the risk associated with family-based applications and broad-based immigration benefit fraud conspiracies; 3) map non-obvious relationships using common data on two or more records to link individuals that would not have been possible otherwise; and 4) to produce a visual representation of relationships for any given individual. While these "relationships" will only be used to discover known risk and fraud patterns, it is important to note that these relationships will need validation by government staff.

In addition the COTS software products facilitate "rule-based analysis" functionality. The result of running link analysis on case data by using the prototype is a list of disclosed and undisclosed/non-obvious relationships and connections. The rule-based analysis tools uses these relationships to create a notification or "flag" of any elements in a subject-based query that reveal connections that a particular risk and fraud rule is designed to reveal. After running the "rule" the applicant's information would reveal such a link and the software will flag that applicant for additional review.



Responsible Officials

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

Approval Signature

Original signed copy on file with the DHS Privacy Office.

Mary Ellen Callahan
Chief Privacy Officer
Department of Homeland Security