Privacy Impact Assessment Update
for the

Eligibility Risk and Fraud Assessment Testing Environment

DHS/USCIS/PIA-29(a)

June 1, 2011

Contact Point
Donald Hawkins
U.S. Citizenship and Immigration Services
(202)272-8000

Reviewing Official
Mary Ellen Callahan
Chief Privacy Officer
Department of Homeland Security
(703) 235-0780

Abstract

The Office of Transformation Coordination of United States Citizenship and Immigration Services is planning to update the Eligibility Risk and Fraud Assessment Testing Environment. This update describes the next phase of this tool to develop, test, and refine the tool's risk and fraud business rules and to load biographic data from legacy systems before deploying to a full production environment.

Introduction

The Department of Homeland Security (DHS) through United States Citizenship and Immigration Services (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 United States. 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 Integrated Operating Environment (IOE) that will make up the transformed environment will include the Eligibility Risk and Fraud Assessment tool to protect national security and the integrity of the legal immigration system. This Privacy Impact Assessment (PIA) Update to the "Eligibility Risk and Fraud Assessment Testing Environment" PIA[1] covers the next phase of testing of the Eligibility Risk and Fraud Assessment tool and the creation of a baseline of data to be used when the IOE is deployed at the end of 2011.

Reason for the PIA Update

The Proof of Concept pilot as detailed in the initial PIA for the Eligibility Risk and Fraud Management tools was successfully able to: 1) match records on a single person from multiple systems to create a single "entity;" 2)identify relationships between individuals based on biographic information; and 3) run potential eligibility risk and fraud rules on the data. At the end of the original Proof of Concept period all of the test data was deleted from the tool. In addition, the Proof of Concept demonstrated the need to begin loading closed and current biographic information from the USCIS legacy systems sooner than expected in order to address the length of time it takes to load the data into the IOE.

The PIA for the Eligibility Risk and Fraud Assessment Testing Environment is being updated for two reasons. First, as part of this previous test USCIS learned that the initial loading of a large amount of

[1] DHS/USCIS/PIA-029 Eligibility Risk and Fraud Assessment Testing Environment (EFRA), April 9, 2010.

data takes a considerable amount of time. The estimates based on the previous load show that the processing time to incorporate legacy data into the tool could range anywhere from 45 to 120 days. It will be necessary to have this data processed and ready for use by the new system the USCIS is preparing to deploy at the end of 2011. This processed data will be loaded into the new tool and used as a baseline for comparing and processing new benefit applications when the tool is deployed.

The second reason for updating this PIA is to continue to refine the capabilities and functionality of the Eligibility Risk and Fraud Management tool. During its initial Proof of Concept, the use of a small subset of closed cases and synthetic data was sufficient to demonstrate that the tool could operate successfully. With this PIA update, USCIS is now beginning to test how the tool will work when it is fully operational by loading all biographic data from selected legacy systems. This is necessary since the underlying software that comprises the Eligibility Risk and Fraud Assessment tool works differently based on the amount and type of data loaded into the tool. Further, the large amount and type of data being loaded is basis for the IOE, so USCIS must be able to identify issues and concerns before it is operational for the public to use. Periodic updates will be provided from the legacy systems to ensure that the baseline of data used in Eligibility Risk and Fraud Assessment tool is timely and accurate when it becomes operational.

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. Given this is a testing environment and the results are not definitive, USCIS will 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 using the legacy systems.

Privacy Impact Analysis

The System and the Information Collected and Stored within the System

During the first test of the Eligibility Risk and Fraud Assessment Testing Environment, the tool utilized a limited amount of closed case data from the legacy USCIS systems and synthetic data for its operations. In this implementation the test tool will utilize Personally Identifiable Information (PII) from both open and closed USCIS case data from the Fraud Detection and National Security Data System (FDNS-DS), 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. Data collected from these legacy systems includes the data elements listed in the initial PIA:

- Name
- Date of Birth
- Place of Birth
- Country of Citizenship
- Gender
- Social Security Number, if applicable
- Alien Number
- Marital Status

- Family Relationships
- Current and Past Address Information
- Current and Past Telephone Information

The following information will also be extracted from the legacy systems:

- Case ID Number
- Application Type
- Passport Information
- Drivers License Number
- Email Address
- Eye Color
- Hair Color
- Height
- Attorney or Accredited Representative Information
- Employment Information
- FBI Number
- Entry/Exit Data

No new information will be collected from individuals for use by this tool.

As noted in the existing PIA, synthetic data from USCIS to simulate misspellings and name variants and from Customs and Border Protection (CBP) to simulate derogatory and border crossing data will be loaded into the Eligibility Risk and Fraud Assessment tool 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.

Uses of the System and the Information

The information will be used to test the ability of the Eligibility Risk and Fraud Assessment tool to: 1) match records on a single person from multiple systems to create a single "entity across multiple systems; 2) establish and maintain links between individuals named on applications or petitions; 3) map non-obvious relationships using common data on two or more records to link individuals; and 4) to produce a visual representation of relationships for any given individual. Part of this process requires testing and validating the process used to load the legacy data into the tool. In addition, USCIS will verify that the "entities" that are created and related to other "entities" by common biographical information, such as address, or passport number, are linked correctly. Also, the biographic information loaded from the legacy systems will be used as a baseline for comparison purposes when the IOE is deployed at the end of 2011.

USCIS has a robust governance process in place including review of rules, policies and procedures. USCIS currently requires potential rules to be developed by a working group of USCIS employees and recommended to a group of delegates from each of the USCIS directorates, called the Program Integrated Product Team (PIPT), the Office of Chief Counsel, and Office of Privacy. With the approval of the PIPT, the rule is presented to the Transformation Leadership Team, consisting of the Director of USCIS and high level staff, for approval before being implemented.

Retention

Although USCIS has not established a retention schedule with NARA for the Eligibility Risk and Fraud Assessment Testing Environment, the original systems of record have retention schedules for the information they retain that have been published and approved by NARA. The Eligibility Risk and Fraud Assessment Testing Environment will retain the information in accordance with those published retention schedules.

Internal Sharing and Disclosure

There are no changes to the internal sharing and disclosures as part of the update to the testing environment.

The information will be used and maintained as part of the testing environment and will not be accessible or shared with any other office within USCIS, or other DHS component agencies, offices or directorates. The information will be used and maintained solely by the OTC to establish the “entity” records, to refine, test, and validate the production tool performance, and to build a baseline of data for the production environment. Ad hoc hits and other results that may identify risk and fraud will not be actionable. However, if the testing environment 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 for further review using the legacy systems.

External Sharing and Disclosure

There are no changes to the external sharing and disclosures as part of the update to the testing environment.

The testing environment will not establish any interfaces with any external organization. The information will be used and maintained solely by the OTC to establish the “entity” records, to refine, test, and validate the production tool performance, and to build a baseline of data for the production environment. Ad hoc hits and other results that may identify risk and fraud will not be actionable. However, if the testing environment 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 for further review using the legacy systems.

Notice

For use in the testing environment, the information is not collected directly from individuals and is used only for internal testing and in preparation of the deployment of the new production tool, 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 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 Update, System of Records Notices (SORN) were published for the applicable systems that are providing the data extracts for the testing environment.

Individual Access, Redress, and Correction

The PII contained within the testing environment is obtained from other USCIS case management systems, 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, as an example it is necessary to file a Change of Address form (AR-11) 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

All privacy risks associated with redress mechanisms are significantly mitigated by the fact that the testing environment exists to assist in assessing privacy risks prior to the deployment of a live production environment. Before the production tool is deployed the necessary PIAs and SORNs will be published detailing the access, redress and correction processes.

Technical Access and Security

There are no changes to the technical access and security procedures as part of the update to the testing environment.

In compliance with federal law and regulations, users who access the testing environment will be granted on a need to know basis. Criteria, procedures, controls, and responsibilities regarding this tool are part of standard operating procedures defining policies and procedures for determining which users may access systems at USCIS. Additionally, there are several department and government-wide regulations and directives, which provide additional guidance and direction.

Technology

There are no changes to the technology as part of the update to the testing environment.

The test environment using the legacy data integrates three COTS software products to perform risk and fraud analysis. This combination of software products: 1) identifies "entities;" 2) establishes and maintains links between individuals named on applications or petitions; 3) maps non-obvious relationships using common data on two or more records to link individuals; and 4) to produces a visual representation of relationships for any given individual. This data will be used during the development, testing, and deployment stages of the DHS System Engineering Lifecycle in establishing the IOE.

Responsible Official

Donald Hawkins, 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