Alien Registration Number Verification via the U.S. Citizenship and Immigration Service's Systematic Alien Verification for Entitlements System

Frances M. Ainslie
Northrop Grumman Technical Services

Kelly R. Buck
Defense Personnel Security Research Center

Alien Registration Number Verification via the U.S. Citizenship and Immigration Service's Systematic Alien Verification for Entitlements System

Frances M. Ainslie, Northrop Gruman Technical Services
Kelly R. Buck, Defense Personnel Security Research Center

Released By – James A. Riedel

BACKGROUND

National Agency Checks in security clearance investigations include checks of alien registration number (ARN) information for foreign-born applicants against the United States Citizenship and Immigration Services Systematic Alien Verification for Entitlements System (SAVE). Currently, these checks are available only on request for most types of investigations. For optimal vetting, all available alien registration information for all foreign-born applicants, and in some cases their spouses and immediate family members, should be checked against SAVE. Additionally, background investigation requirements for government credentialing will require higher volume use of SAVE for permanent foreign residents in the United States. The purpose of this study was to evaluate the implications of conducting high-volume automated checks of SAVE to vet ARNs of applicants for security clearances and government credentials.

HIGHLIGHTS

Of the 9,983 subjects submitted for ARN verification, 4% or 417 could not be verified. The majority of the unverified subjects were military applicants (n=323 or 77.4% of unverified subjects). While the majority of investigation outcomes for unverified subjects were classified as "unknown," 89 subjects were adjudicated as eligible for access to classified information. Reasons for discrepancies in the ARN verifications could be explained for 53 of these subjects, but 31 cases did not contain comments that would allow understanding of the differing results. Two of the 89 unverified subjects' case summaries were too sparse to classify in any way; 8 had an unfavorable INS check or no record was found, neither of which was elaborated; and 26 appeared to have had no INS check. Many subjects with unverified ARNs report the United States as their place of birth. Findings demonstrate the importance of fully resolving legal immigration status in background investigations.

REPORT DOCUMENTATION PAGE

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE: 03-01-2008	2. REPORT TYPE Technical Report 08-01	3. DATES COVERED (From – To) January 2006 - January 2008

4. Alien Registration Number Verification via the U.S. Citizenship and Immigration Service's Systematic Alien Verification for Entitlements System	5a. CONTRACT NUMBER:
	5b. GRANT NUMBER:
	5c. PROGRAM ELEMENT NUMBER:

6. AUTHOR(S) Ainslie, Frances M. Buck, Kelly R.	5d. PROJECT NUMBER: SS-06-04
	5e. TASK NUMBER:
	5f. WORK UNIT NUMBER:

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Defense Personnel Security Research Center 99 Pacific Street, Suite 455-E Monterey, CA 93940-2497	8. PERFORMING ORGANIZATION REPORT NUMBER PERSEREC: Technical Report 08-01
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Defense Personnel Security Research Center 99 Pacific Street, Suite 455-E Monterey, CA 93940-2497	10. SPONSORING/MONITOR'S ACRONYM(S): P&R, CI&S CCF
	11. SPONSORING/MONITOR'S REPORT NUMBER(S):

12. DISTRIBUTION/AVAILABILITY STATEMENT: Distribution Unlimited

13. SUPPLEMENTARY NOTES:

14. ABSTRACT: The purpose of this study was to evaluate the implications of conducting high-volume automated checks of the United States Citizenship and Immigration Services Systematic Alien Verification for Entitlements System (SAVE) to vet alien registration numbers of applicants for security clearances and government credentials. Of the 9,983 subjects submitted for alien registration number verification, 4% or 417 could not be verified. The majority of the unverified subjects were military applicants (n=323 or 77.4% of unverified subjects). While the majority of investigation outcomes for unverified subjects were classified as "unknown," 89 subjects were adjudicated as eligible for access to classified information. Reasons for discrepancies in the ARN verifications could be explained for 53 of these subjects, but 31 cases did not contain comments that would allow understanding of the differing results. Two of the 89 unverified subjects' case summaries were too sparse to classify in any way; 8 had an unfavorable INS check or no record was found, neither of which was elaborated; and 26 appeared to have had no INS check. Many subjects with unverified ARNs reported the United States as their place of birth. Findings point out the importance of fully resolving legal immigration status in background investigations. Recommendations for resolving discrepancies are provided.

SUBJECT TERMS: Identity vetting, Registered alien, Alien Registration Number, ARN, Alien Status Verification, Citizenship Verification, USCIS SAVE, Data Quality, Automated checks, Automated record checks, Investigative requirements, INS checks, Identity fraud, Falsification

15. SECURITY CLASSIFICATION OF: UNCLASSIFIED			16. LIMITATION OF ABSTRACT:	17. NUMBER OF PAGES: 41	19a. NAME OF RESPONSIBLE PERSON: James A. Riedel, Director
a. REPORT: UNCLASSIFIED	b. ABSTRACT: UNCLASSIFIED	c. THIS PAGE: UNCLASSIFIED			19b. TELEPHONE NUMBER (Include area code): 831-657-3000

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI td. Z39.18

PREFACE

This study is part of ongoing efforts by PERSEREC to identify electronic data resources that support high-volume, cost-effective, automated vetting of personnel for government credentials and eligibility for access to classified information. The results in this report pertain to vetting of foreign-born applicants who request access to sensitive information, facilities, and duties. National policy and procedures should promote maximum vetting of all foreign-born individuals' legal immigration or residency status in the United States before granting access to federally controlled information, facilities, and material.

James A. Riedel
Director

PREFACE

EXECUTIVE SUMMARY

INTRODUCTION

Current OPM policy for suitability and SECRET-level security clearance investigations is to only validate alien registration number (ARN) information if requested by offices submitting applications for background investigations. This practice requires tens of thousands of human resource and personnel security specialists to recognize the need for conducting ARN checks, to know how to request the checks, and perhaps to know how to resolve discrepant information when found. The purpose of this study was to evaluate the utility and implications of systematically conducting high-volume automated checks of the United States Citizenship and Immigration Services (USCISC) Systematic Alien Verification for Entitlements System (SAVE) to vet ARNs of all applicants for security clearances or government credentials who are not U.S. citizens by birth.

METHODOLOGY

A random sample of 10,000 subjects was taken from the Defense Security Service Case Control Management System (CCMS) and the Joint Personnel Adjudication System (JPAS). The criteria for inclusion were subjects who reported an ARN and who had valid authorizations for release of information.

Data submitted to USCIS for verification included ARN, last name, first name, gender, DOB, and Document Type. Data were submitted, in a single batch file, to USCIS in a format specified in their Customer Processing System (CPS) Automated Secondary Batch Access Method Interface Control Document Agency Version, Draft v38. File submission and retrieval were accomplished via secure file transfer protocol (SFTP).

RESULTS

Of the 9,983 subjects processed by USCIS for ARN verification, 4% or 417 could not be verified. The majority of the unverified subjects were military applicants (n=323 or 77.4% of unverified subjects). While the majority of investigation outcomes for unverified subjects were classified as "unknown," 89 subjects were adjudicated as eligible for access to classified information. Reasons for discrepancies in the ARN verifications could be explained for 53 of these subjects, but 31 cases did not contain comments that would allow understanding of the discrepancies. Two of the 89 unverified subjects' case summaries were too sparse to classify in any way; 8 had an unfavorable INS check or no record was found, neither of which was elaborated. Finally, 26 subjects appeared to have had no INS check. Many subjects with unverified ARNs reported the United States as their place of birth. Just over 6% of subjects were listed as verified by USCIS but had significant data disparities between that submitted by DoD and that on file in SAVE. These discrepancies may provide valuable data for investigating unreported aliases and other falsification.

RECOMMENDATIONS

Improve the Verification of Citizenship and Alien Data

Current policy for vetting ARN information places full responsibility on front-line personnel to know when and how to request validation. It is at their discretion to rely solely on the information provided by applicants as recorded on documents that the applicants present vice corroborating the information through independent sources. If applicants have become naturalized since immigrating to the United States, personnel security policy and procedures do not even require recording of ARNs that were issued at the time of immigration. Current policies and procedures reflect the importance placed by the personnel security system on documenting that subjects meet citizenship requirements but do little in the way of protecting against identity and application fraud on the part of those same subjects. In light of the availability of independent corroboration of immigration and naturalization information through USCIS, better use of this information should be required.

- At a minimum, subjects born outside the United States to other than U.S. parents should be required to provide their ARN, even if they have subsequently become naturalized U.S. citizens.

- Use the SAVE system to validate ARN information for all personnel who have been issued ARNs, regardless of their citizenship status at the time of investigation.

- Where applicants have gained entry through granting of asylum, review their paperwork for requesting asylum.

- Adopt investigative guidelines for case expansion of investigations with unverified ARN information or verified ARN information with significant discrepancies. Ensure aliases that develop are documented and included in subsequent components of investigations.

- Where applicants are verified but with significant name discrepancies, determine why names were listed as they were and ensure developed aliases are included in other record checks.

- Because reporting the United States as the place of birth (POB) appears to be a relatively common data error, ensure that the true POB is obtained and correctly documented. Subjects reporting non-U.S. citizenship, dual citizenship, or foreign passports (current or expired) should be questioned as to their POB or citizenship status.

- Human Resource and Security officers should be trained and required to review documents, and review relevant portions of personnel security questionnaires and attest to their completeness and accuracy. As is done at the MEPS, background investigations should not proceed until ARNs are supplied.

Additionally, security questionnaires should include a question for aliens regarding the name in which their ARNs were issued. One would assume that the 'original' name would be reported in the Alias section of the form. However, multiple aliases

may exist, which would require multiple USCIS queries. An item specifically addressing the name associated with the ARN would allow investigators to check USCIS once, with the correct name. PERSEREC has made a similar recommendation for SSNs in the study addressing the verification of Social Security Numbers (SSN) via the Social Security Administration's Enumeration Verification System. This approach has been adopted by the Internal Revenue Service (IRS) instructions for electronic filers. The IRS instructs filers to report their names *exactly* as they appear on their Social Security Cards, regardless of the names on their birth certificates.

Finally, while the USCIS Secondary Verification process was not designed to provide identifying information, modifying the response to include the same data that are returned in the Primary Verification would aid in identifying errant ARNs, data discrepancies and aliases, and support the documentation needs of the personnel security program.

Develop and Implement Investigative Procedures When ARNs are not Verified

If subjects cannot be verified, or they have been verified but significant discrepancies exist between SAVE and reported data, the subjects should be interviewed to rule out errant data or to obtain the identifying data with which they entered the United States. If aliens cannot be verified via SAVE, then their investigations should not conclude until resolved.

Subject interviews to resolve ARN discrepancies should consist of reviewing applicants' Alien Registration Cards or Citizenship or Naturalization[1] Papers. Data on source documents should be compared to data that were sent for verification. If errors are with the data provided to USCIS, then they should be corrected and resubmitted.

If, on the other hand, applicants claim that the submitted data were correct, then they should be directed to the USCIS website where they can contact USCIS or find the nearest USCIS Field Office to resolve their cases.

Possible new investigative standards upon a finding of discrepancies between USCIS files and information provided by subjects are as follows:

- Request subjects to show their official Alien Registration Card to determine if discrepancies are due to data entry errors.

- If discrepancies cannot be resolved favorably for subjects, suspend investigations.

- Document reasons for suspension of investigations in JPAS, the SII, or other appropriate systems.

- Instruct subjects to resolve issues with USCIS.

---

[1] The ARN appears on the Naturalization Certificate, but is identified as either 'INS Registration Number' or 'CIS Registration Number.'

- Resume processing only after subjects provide evidence of resolution and ARN can be verified.

- If discrepancies cannot be resolved through USCIS or if subjects fail to follow through with efforts to resolve discrepancies, notify appropriate authorities[2] and document status in JPAS.

- Identify appropriate authorities for resolving potentially fraudulent ARNs.

For cases that cannot be resolved as data entry errors on the part of DoD or USCIS, DUSD(CI&S) will need to develop policy for investigation, adjudication, and flagging those subjects in the event they try to reapply at a later time.

Develop and Implement Investigative Procedures When ARNs are Verified With Discrepancies

For a given ARN, SAVE checks require only matching the month and day of the date of birth and the first letter of the first name. Consequently, SAVE may uncover days of birth, unreported aliases, and countries of birth that differ from those reported by subjects. Therefore:

- Verified responses should be reviewed by investigators or adjudicators to ensure that data on file for immigrant subjects in DoD are consistent with those on record with USCIS.

- When discrepancies are deemed significant, they should trigger expanded investigation and special interviews with subjects at the same level as what would be required when subjects are reported as unverified by USCIS.

- Given the importance of identity vetting and the ability to verify an individual across widely varying systems, it is worth considering a modification to the manner in which aliases are stored. If an alias has been used to verify alien status, then that alias should be so marked. Likewise, if an alias (or maiden name) was used to verify an SSN, it should be so marked. This would preclude problems with later verification attempts and provide the documentation necessary to support quality personnel security investigations and adjudications.

Improve Data Quality

The Government Accountability Office (GAO-07-310, 2007) identified four actions that DoD must take to improve its personnel security clearance program. One of these actions was to implement procedures to eliminate documentation problems. Providing a mechanism to correct errant data would greatly improve electronic documentation and could result in fewer requests for supporting documentation.

Review Data Holdings. The growing pervasiveness of database sources that can be used for research and management decisionmaking means that more attention must be focused on data quality. Researchers must become intimately familiar with

---

[2] Exactly who will be designated as appropriate authorities to deal with suspected identity fraud will need to be determined by the DUSD(CI&S) Director of Security.

the data sources in order to correctly interpret anomalous results. Since the quality of documentation differs dramatically from one source to the next, it is imperative that users obtain data samples and review them in light of matching requirements and the provided documentation. Discrepancies should be discussed with the data holder, in order to reach an understanding of those discrepancies and deal with them appropriately. The importance of reviewing and analyzing the data cannot be overstated.

Implement Quality Controls. Data quality problems can be prevented with data entry controls. When entering data using an automated tool, users should be allowed to enter unconstrained data only where absolutely necessary (SSN, ARN, name, DOB, addresses, etc.). While the content of these fields cannot be constrained, the length of SSN, ARN, and DOB can, ensuring at least this degree of conformance. Additionally, critical data elements, such as SSN and ARN, should require double entry to prevent errors.

For coded values, users should be provided lists of allowed values from which to choose. And, when the relationship between two fields is unambiguous, data entry should be so constrained. For example, users who indicate they were not born in the United States should be guided to the entry of an ARN. Likewise, users who enter ARNs and indicate that they were born in the United States should be instructed to review the conflicting entries.

Until cross-referencing of related data elements can be automated security officers should fill the gap with careful review of related fields and increased emphasis on immigration data. Sponsoring security officers should implement a program to review all data entry and obtain ARNs before submitting paperwork for the non-U.S. citizens.

Future Directions

DoD should also develop policy and procedures for vetting ARN information provided by applicants against ARN information already on file in DoD systems. Doing so would defend against subjects borrowing each other's personal identifiers or using the same criminal sources that sell fraudulent identifications to willing buyers.

EXECUTIVE SUMMARY

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The Defense Security Service (DSS) tasked the Defense Personnel Security Research Center (PERSEREC) with evaluating the feasibility of conducting high-volume checks of the United States Citizenship and Immigration Services (USCIS) Systematic Alien Verification for Entitlements System (SAVE). SAVE is a system that processes private employers' and government agencies' requests for verification of foreign-born personnel's legal immigration and citizenship status. From an employment standpoint, the purpose of the verification is to ensure that U.S. citizens and noncitizens are eligible for employment. From a security standpoint, SAVE provides a means for ensuring that foreign-born applicants are who they say they are and have entered the country through legal channels under names that are known to the DoD.

The Standard Form 86, *Questionnaire for National Security Positions* (SF-86) is the means for all military entrants and subjects of national security clearances investigations to report their citizenship status. If they indicate they are not U.S. citizens, then they are asked to provide their Alien Registration Numbers (ARN) and countries of citizenship. Additionally, if their fathers, sisters, brothers, children, current spouses, or persons with whom they have spouse-like relationships are not U.S. citizens, they are asked to provide their relatives' and associates' ARNs.

The SF-86 does not request, however, ARNs of immigrants who became U.S. citizens through naturalization. Instead, the form requests only naturalization certificate numbers. Future research will examine the utility of using the SAVE system to conduct high-volume validation of naturalization numbers.

For those subjects who do provide ARNs for themselves and their immediate family members, the SF-86 instructs them to indicate the date and place where they entered the United States. Although Section 15 of the form (where relatives'/associations' citizenship information is entered) instructs the applicant to provide ARNs and Place of Entry for alien relatives, this information must be entered in the free-form 'Additional Information' block. As applicants rely on forms to guide them in their responses (irrespective of instructions), it is likely that this leads to suboptimal reporting of the relatives' ARNs. And, for checking ARNs against SAVE, only ARN, last name, first name, DOB, gender, and the document type that was offered as proof of ARN are submitted.

DoD contracts all of its investigations to the Office of Personnel Management (OPM). OPM's current policy is to only validate ARN information if requested by the security officer responsible for submitting the investigation. DoD philosophy, however, is to validate ARN information for all subjects and on all relatives and associates for whom ARNs are available. As such, DoD would be interested in higher-volume checks than are being conducted based only on individual requests.

The above descriptions indicate that DoD could make much greater use of SAVE for higher-volume ARN validation checks. This study examines the feasibility for doing

INTRODUCTION

so, determines the proportions of ARNs that can be validated, and the rates of problematic matching results. These findings are useful for understanding the time and resource implications of greater use of SAVE for validating immigration, employment, and legal residence status of military and security clearance applicants.

To date, DoD and OPM do not have statistically representative data on probable results from high-volume submissions of ARN information to the USCIS SAVE system. Additionally, where data cannot be validated, information is lacking on the reasons why. Possibilities include errant or fraudulent information being provided by applicants, data entry errors by the applicant, DoD or OPM personnel while processing the information, or data quality errors on the part of USCIS. Findings from this study will clarify the feasibility of validating ARN information in light of these possible outcomes.

As ARNs are validated in higher numbers, investigators and adjudicators will have to address more cases with problematic matching results. As such, this report provides preliminary recommendations on steps to take to respond to, and where possible, resolve problematic matches.

METHODOLOGY

SUBJECTS

A random sample of 10,000 subjects was taken from the Defense Security Service Case Control Management System (CCMS) and the Joint Personnel Adjudication System (JPAS). The criteria for inclusion were subjects who reported an Alien Registration Number (ARN) who had valid authorizations for release of information. That is, to be included in the study, subjects had to have an ARN and had a JPAS status of Active or Reactivated to indicate that they had not terminated their affiliation with DoD and had signed their SF-85/SF-86 information release within the last 5 years. The number of subjects who met these criteria was 13,508. Using the SPSS random selection function, the sample of 10,000 subjects was extracted.

All data required for this study were pulled from JPAS and CCMS. CCMS contained self-reported data from subjects' *Questionnaire for National Security Positions* (SF-86) or other relevant form (e.g., SF-85, *Questionnaire for Nonsensitive Positions*). JPAS contained investigator-reported information regarding the status and determination of an investigation.

Table 1 identifies the sample characteristics and provides the percentage of the population who possessed each characteristic. The sample and population percentages were very similar for each characteristic, indicating that the sample was representative.

THE SOURCE DATABASES

Joint Personnel Adjudication System (JPAS)

JPAS is the centralized database that provides real-time information regarding clearance, access, and investigative status of individuals in the DoD Personnel Security Program. It supports central adjudication facility (CAF) functionality by providing an information system to assist the adjudication process and provides security managers with a means of checking the security clearance status of individuals in the DoD Personnel Security Program.

JPAS contains a record of each person for whom a security investigation has been performed. The relevant data for this study extracted from JPAS for each subject were:

- Subject status with regard to the position for which investigations were conducted (Active, Reactivated, or Archived)
- Eligibility[3] granted or not granted (e.g., denied, revoked, Secret, Top Secret, etc.)
- Investigator comments, interviews, and summaries

---

[3] The highest level of classified information which may be disclosed to a person based on the type of completed investigation.

Table 1
Sample Characteristics (n = 9983)

		Sample n	% of Sample[4]	% of Population[5]
Gender	Female	2,341	23.4	23.1
	Male	7,629	76.4	76.9
	Unknown	13	.1	.1
Age	< 19	3	.0	.0
	19 - 21	230	2.3	2.2
	22 - 24	1,926	19.3	17.5
	25 - 29	5,028	50.4	50.8
	30 - 34	1,381	13.8	13.7
	35 - 39	731	7.3	7.2
	40 - 44	351	3.5	3.6
	45 - 50	203	2.0	1.9
	> 50	130	1.3	1.4
	Unknown	0	.0	1.8
Investigation Type	NACLC	2,606	26.1	21
	NACLC-PR	33	.3	.4
	SSBI	141	1.4	1.6
	SSBI-PR	19	.2	.2
	Other[6]	6,930	69.4	67.5
	Unknown	254	2.5	9.3
Job Status	Civilian	132	1.3	1.3
	Contractor	249	2.5	2.5
	Military	4,839	48.5	55.9
	Unknown	4,763	47.7	40.3
Citizenship	U.S. Citizen	3,909	39.2	38.6
	Not a U.S. Citizen	4,963	49.7	50.3
	Dual Citizen	6	.1	.0
	Unknown	1,105	11.1	11.1

Defense Security Service Case Control Management System (CCMS)

CCMS is a recently retired automated Defense Security Service (DSS) information system that was used for opening, tracking, and closing personnel security investigation cases. It contained the responses to applicants' relevant security questionnaires (e.g., SF-85, SF-86), organized in relational database tables. The data extracted from CCMS for each subject were:

---

[4] Percentages may not sum to 100, due to rounding.
[5] Percentages may not sum to 100, due to rounding.
[6] The Investigation Types that are included in the 'Other' category are identified in Table 4.

- All reported names (first, middle, last, maiden, alias, etc.)

- Date of Birth (DOB)

- Gender

- Citizenship

- Job Status (Civilian, Military, Industry)

- Country of Birth

- Alien Registration Number (ARN)

- Investigation Type

Systematic Alien Verification for Entitlements System (SAVE)

SAVE incorporates two levels of verification. The first, Primary Verification, is an automated check of USCIS computer records. If the results of that check are insufficient (ARN is not found or identifying information is significantly different), then a Secondary Verification is performed. The Secondary Verification involves a manual review of SAVE and other USCIS information systems, including paper copies of relevant documents.

The data submitted to SAVE are ARN, last name, first name, DOB, gender, and the document type that was offered as proof of ARN. If the document type is not found on the SAVE list of accepted documents, then additional information regarding the document (document type, issue date, expiration date, etc.) is also submitted. These data elements are reported on the SF-86.

USCIS requires that, in the submission file, the ARN be stripped of the leading 'A[7] and any other letters or extraneous characters (spaces, dashes, etc.). This is the only data modification required in the submission file.

In the Primary Verification – which is automated – the following data elements must match SAVE: ARN, first character of the first name, and the month and year of DOB. If these elements do not match, then the transaction is sent to Secondary Verification, where the match criteria are unbounded, as the match is performed by status verifiers (that is, by persons and not by computer).

USCIS returns the alien/employment status of the individual, at time of inquiry. In the Primary Verification Response, additional data elements – as they appear in SAVE – are also returned, such as last name, first name, DOB, and place of birth (POB). These data enable comparing self-reported information with the information in SAVE, and identifying discrepancies. These data elements are not returned in the Secondary Verification Response; only the alien employment status is returned.

---

[7] The US Citizenship and Immigration Service (USCIS) confirmed that all valid ARNs begin with 'A.'

Thus, data discrepancies cannot be investigated for subjects undergoing Secondary Verification.

ARN ANOMALIES

Once the sample had been drawn, the ARN was cleaned; letters and extraneous characters were removed. After cleaning, several subjects were left with a blank ARN. Review of the reported ARNs revealed entries such as 'Lost it,' 'Applied for,' 'Antilles,' 'A,' and so on. Thus, a new sample was drawn, based on the presence of an ARN after cleaning.

It should be noted that the initial subject selection was based on a reported place of birth outside the United States. However, the number of these subjects reporting an ARN was so low that the selection criterion was changed to the presence of an ARN, which is required for matching to SAVE.

DATA SUBMISSION

Data submitted to USCIS for verification were ARN, last name, first name, gender, DOB, and Document Type. The data were submitted, in a single batch file, to USCIS in their standard format, as identified in their Customer Processing System (CPS) Automated Secondary Batch Access Method Interface Control Document Agency Version, Draft v38 (U.S. Department of Homeland Security, 2006). File submission and retrieval were accomplished via secure file transfer protocol (SFTP). All data were submitted as they appeared in CCMS, with the exception of ARN, as described above. Figure 1 presents the overall Verification Process in a graphical manner.

Primary Verification

In the Primary Verification Response, USCIS returned the submitted ARN, as well as identifying data from SAVE and other administrative data elements. Also returned was an indication that a subject was sent to Secondary Verification (to be retrieved at a later date). If the case was verified during Primary Verification, then a status code was returned with one of the following values:

- Lawful Permanent Resident-Employment Authorized
- Cuban/Haitian Entrant-Temporary Employment Authorized
- Institute Additional Verification
- Refugee - Employment Authorized
- Asylee - Employment Authorized
- Temporary Resident - Temporary Employment Authorized
- United States Citizen

The identifying data (last name, first name, date of birth, and place of birth) returned in the Primary Verification supported a comparison of reported data and those in SAVE, for verified subjects. The comparison yielded discrepancies, which are presented along with the verification results.

USCIS VERIFICATION PROCESS

PERSEREC	USCIS

Inquiry

Request Verification

→ ARN
Last Name
First Name
Gender
DOB
Document Info →

Process against SAVE

Primary Verifications

1. Record Alien Status
2. Identify Discrepancies in Last Name, First Name, DOB, POB
3. Record Discrepancies

← ARN
Last Name
First Name
DOB
POB
Alien Status

Matched in SAVE

Retrieve Secondary Verification

← Retrieval Identifier

Not Matched in SAVE

Secondary Verifications

Retrieval Identifier

Record Alien Status

Verify against Alternate Sources

Alien Status

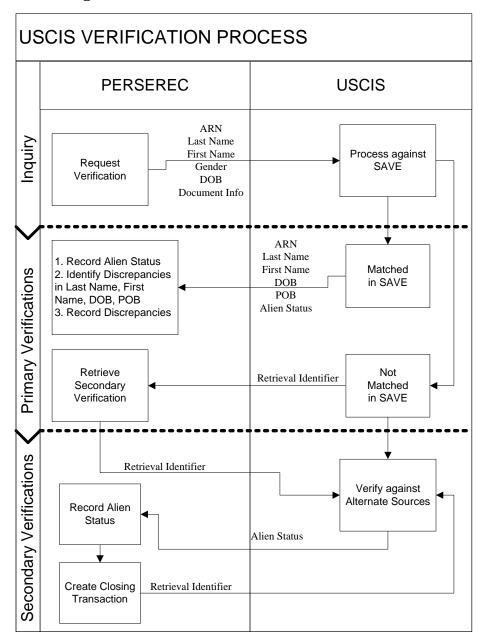Create Closing Transaction

Retrieval Identifier

Figure 1 USCIS Verification Process

Secondary Verification

If a subject required Secondary Verification, then the response had to be retrieved from USCIS, with a different transaction format using only the Retrieval Identifier

returned by USCIS, after a minimum of 72 hours. The data returned in this response included an expanded list of status codes (as well as administrative data elements). The alien status codes were mapped to those returned in the Primary Verification, for ease of reporting. Three values could not be mapped to the primary verification set and thus were added to the response set:

- Resubmit Doc (need copy original)
- Conditional Resident, Employment Authorized
- Parolee, Employment not Authorized.

Identifying data from SAVE were not returned in the Secondary Verification Response. Thus, the discrepancy analysis could not be performed on these subjects. After the retrieval of the Secondary Response, USCIS required a final transaction indicating that the case was closed. A closing transaction was submitted for all subjects who underwent secondary verification.

RESULTS

OVERALL RESULTS OF ARN VERIFICATION

As discussed, 10,000 subjects were submitted for verification. Of these, 16 did not have a date of birth and thus were not processed by USCIS; additionally, there was one duplicate in the file. Thus, results are based on 9,983 subjects. Of these, 86.3% (8,617) were verified in the Primary Verification Response. The remaining subjects (1,366) were sent to Secondary Verification, where 949 were verified and documentation was requested for 417.

The final verification values returned by USCIS (after secondary verifications were completed) are provided in Table 2. The shaded area provides the detailed status of verified subjects. Nearly 96% of subjects were verified and authorized employment; only two verified subjects were not authorized employment.[8]

One interesting finding is the 636 subjects were verified through SAVE but SAVE had information different from that provided by applicants. These differences were identified by comparing the reported data with that returned by USCIS in the Primary Verification Response. In some cases, the results of that comparison yield unreported aliases, as will be discussed in a later section.

USCIS requested hard copy documentation for just over 4% of subjects. These 417 subjects were subjected to secondary verification. The requests for documentation indicate that USCIS could not validate DoD data using SAVE or other information available during the Secondary Verification Process.

Table 2
Overall Results of USCIS Verification (n = 9,983)

Alien Status	n	%[9]
Verified: Authorized Employment, no Data Discrepancies	8,941	89.6
Verified: Not Authorized Employment, no Data Discrepancies	2	.0
Verified: Authorized Employment, with Data Discrepancies	623	6.2
Total Verified	9566	95.8
Not Verified: original copy of document requested	417	4.2
U.S. Citizen	6,333	63.4
Lawful permanent resident	3,222	32.3
Asylee/refugee, employment authorized	3	.0
Asylee – employment authorized	2	.0
Conditional resident, employment authorized	1	.0
Temporary resident – temporary employment authorized	3	.0
Parolee, employment not authorized	2	.0

---

[8] Parolees must apply for and be approved for employment authorization. If the authorization is granted, the status changes to the appropriate authorized condition.
[9] Percentages may not sum to 100, due to rounding.

RESULTS

<div align="center">

Table 3
Characteristics of Unverified Subjects (n = 417)

		n	%[10]
Gender	Female	98	23.5
	Male	318	76.3
	Unknown	1	.2
Age	< 19	0	--
	19 - 21	14	3.4
	22 - 24	86	20.6
	25 - 29	194	46.5
	30 - 34	51	12.2
	35 - 39	32	7.7
	40 - 44	14	3.4
	45 - 50	17	4.1
	> 50	9	2.2
Investigation Type	NACLC	122	29.3
	NACLC-PR	2	.5
	SSBI	8	1.9
	SSBI-PR	4	1.0
	Other	265	63.5
	Unknown	16	3.8
Job Status	Civilian	4	1.0
	Contractor	18	4.3
	Military	323	77.4
	Unknown	72[11]	17.3
Citizenship	U.S. Citizen	165	39.6
	Not a U.S. Citizen	196	47
	Dual Citizen	0	--
	Unknown	56	13.4

</div>

The majority of subjects fell into the 'Other' Investigation Type. This category was expanded to examine investigation outcomes as shown in Table 4. Of those with 'Other' Investigation Type, 122 subjects underwent an ENTNAC investigation for military entrance which enabled us to classify them as Military if their job status was unknown. Thus, the majority of unverified subjects were military applicants (323 subjects, or 77.4%).

Table 4 provides the investigation outcome/eligibility for the unverified subjects. While Table 4 presents the expanded list of Investigation Types, all NACLC investigations were combined into a single category for ease of presentation.

---

[10] Percentages may not sum to 100, due to rounding.
[11] Military job status was inferred from the ENTNAC investigation type for 122 subjects. The job status for these 122 subjects was 'Unknown' in JPAS/CCMS.

Outcomes were available for 123 investigations, 89 of which were adjudicated favorably or were granted a clearance – the shaded area in Table 4.

Table 4
Investigation Outcomes for Unverified Subjects (n = 417)

	Investigation Type								
Eligibility	ENTNAC	NAC	NACLC	SAC	SSBI-PR	SSBI	XNAC	Un-known	Total
Favorable	4	2	9	2	0	0	0	1	18
Interim Secret	0	0	4	0	0	0	0	0	4
Interim TS	1	0	0	0	0	0	0	0	1
LAA Secret	0	0	1	1	0	0	0	0	2
SCI	2	0	0	0	3	4	0	3	12
Secret	7	1	41	0	0	0	1	1	51
Top Secret	0	0	0	0	0	1	0	0	1
Action Pending	2	1	0	0	0	0	0	0	3
Revoked	0	0	0	0	0	1	0	0	1
No Determination	8	3	13	1	0	0	0	0	25
Loss of Jurisdiction	2	1	1	0	0	0	0	1	5
Unknown	162	60	55	2	1	2	2	10	294
Total	188	68	124	6	4	8	3	16	417

UNVERIFIED SUBJECTS WITH FAVORABLE ADJUDICATIONS

To further evaluate the unverified subjects with favorable adjudications, PERSEREC researched additional information on these subjects in CCMS and JPAS – in particular investigator comments and results of interviews. We found additional information identified in the investigation, such as a corrected ARN or developed alias, documented in these comment fields. Table 5 provides the results of this review. The review identified favorable INS[12] Checks for 53 subjects, leaving only 36 subjects unverified by USCIS but with favorable investigative outcomes.

Table 5
INS Checks: Unverified Subjects with Favorable Investigative Outcomes (n = 89)

Investigative Comment Classification	n
Favorable INS Check with corrected ARN (typo in reported ARN)	11
Favorable INS Check with same ARN and developed or reported alias	11
Favorable INS Check with no additional information	31
Total Favorable INS Checks Reported	53
Unfavorable INS check or INS record not found	8
No INS check reported	26
Insufficient information to further classify	2
Total Unfavorable/Unexplained INS Checks	36

[12] USCIS was previously known as the Immigration and Naturalization Service (INS).

RESULTS

Of the 89 unverified subjects adjudicated favorably, 53 subjects had favorable INS checks recorded in their investigations. Eleven of these subjects were verified with an ARN that differs from that in CCMS and which appeared to be typographical errors. In most cases the corrected ARNs were found in the INS check remarks, but others were found in the subject interview summary or other comment field. Another 11 subjects were verified with the same ARN and a developed or reported alias. The remaining 31 cases did not contain comments that would allow us to further classify them in a meaningful manner.

Two of the 89 unverified subjects' case summaries were too sparse to classify in any way; 8 had an unfavorable INS check or no record was found, neither of which was elaborated; and 26 appeared to have had no INS check. Many, but not all, of this latter group reported the United States as their place of birth. However, many of the subjects for whom an INS check was conducted also reported the United States as their place of birth. This seemed to be a fairly common data error, as 346 of all subjects reported the United States as their place of birth, even though they had reported an ARN. Of course, the possibility exists that, for some subjects, the ARN is reported in error, and the place of birth is indeed the United States.

Inspection of the comments in these case summaries revealed a shortcoming in relying on coded data for individual cases. Since JPAS and CCMS have no mechanism for data correction, if one is investigating a particular case, then all documentation must be reviewed. In some cases, even case summaries will not provide the relevant information. In the following example two subjects reported the same ARN and were verified in the Secondary Verification response. Because no identifying data were returned (including the ARN) in this response, the data did not appear to support the verifications.

> Two subjects reported the same ARN; they had the same last name, middle name, and gender, but a different SSN, date of birth, and first name. In the Primary Response, USCIS returned the name, date of birth, and gender of the individual to whom the reported ARN belonged. This individual was the mother of each subject, who were brothers. In the Secondary Response the subjects were verified.

> A review of case summaries found reference to favorable INS checks with developed aliases, but not to a corrected ARN. We contacted USCIS regarding these subjects. Using the mother's ARN and the subjects' names, as was done during the Secondary Verification, they provided us with the each subject's ARN.

Because no identifying data are returned in the Secondary Response, the investigators did not know that a different ARN was used for verification and thus made no note of it. What appeared to be fraudulent use of identity turned out to be a lack of available data.

Use of data holdings such as JPAS can provide relevant and meaningful information regarding the Personnel Security Program population. However, much like theoretical or behavioral research, where one attempts to predict the behavior of a certain percentage of a population, but never the behavior of an individual, inferences about a given subject based on coded data alone may prove problematic. It should be noted that, while these systems were not designed for research purposes, the degree to which they support it is impressive. However, as they are used more and more for this purpose – and for administrative decisionmaking – additional focus should be placed on ensuring their accuracy and completeness.

UNVERIFIED SUBJECTS WITH UNKNOWN ADJUDICATIONS

The number of unverified subjects with an unknown adjudication was 294. Of these, 162 underwent an ENTNAC investigation, the outcome of which would not be a security clearance – although one might expect that the investigation outcome would be recorded as Favorable/Unfavorable. While it is doubtful that any subject would be granted access to classified information without an eligibility code present in JPAS, 13 of these subjects were granted access. Nine subjects were granted Interim Secret access, 3 were granted Secret access, and 1 was granted Confidential access.

While these subjects would not be granted access to classified information, the fact remains that their status cannot be verified with USCIS. Further research on this group revealed that 15 subjects have separated since their investigation (based on person status code in JPAS). Investigative summaries were reviewed for information regarding INS checks for the remaining 279 subjects, and are presented in Table 6.

Like the unverified subjects who had favorable investigative outcomes, many of the supporting alien verification data can only be found in investigative comments and are not reflected in the reported data. A total of 211 subjects had favorable INS responses; 118 of these had modifications to reported data. Ninety-three subjects had favorable responses with no reference to data modifications. However, given that identifying data are not returned in a Secondary Verification Response, it is possible that the investigator was not aware that submitted data were not sufficient to verify the subject.

Table 6
INS Checks: Unverified Subjects with Unknown Investigative Outcomes (n = 279)

Reason for Initial Unverified Status	Count
Favorable INS Check with some modifications	**118**
• same ARN & alias	8
• same ARN & developed alias	19
• same ARN developed alias & different DOB	1
• different ARN	70
• different ARN; initial response, no record on file	1
• different ARN & alias	4
• different ARN and developed alias	10
• different ARN & developed alias; initial response unfavorable	1
• different ARN and developed alias & different DOB	1
• different ARN & DOB	1
• no referenced ARN & developed alias	2
Favorable INS Check with no modifications referenced	**93**
• same ARN	78
• same ARN; initial response unfavorable	2
• no referenced ARN	5
• no referenced ARN; initial response unfavorable	1
• no referenced ARN; initial response unfavorable with different ARN	1
• no referenced ARN; initial response unfavorable with same ARN	1
• no referenced ARN; initial response no record on file	4
• no referenced ARN; initial response insufficient data	1
No INS Record on File	**25**
No Investigative Report Available	**7**
No INS Check Referenced	**28**
• POB = U.S.	16
• Skeletal Record	5
• No comment	7
Unfavorable INS Check	**6**
• Same ARN	2
• Different ARN	1
• No referenced ARN	2
• No referenced ARN; initial response favorable with different ARN	1
No INS Response	**2**

No INS record was found for 25 subjects and another 6 received an unfavorable INS check. With the current emphasis on identity vetting, PERSEREC suggests that these 31 investigations should have been halted and returned to the sponsoring agency to review the data with the subject and make corrections, as necessary.

DISCREPANCIES BETWEEN REPORTED AND USCIS DATA

The discrepancy analysis can be conducted only on subjects who were verified in the Primary Verification Response, as the required data are returned only in this response. The first name, last name, DOB, and POB returned by USCIS were

compared with associated data in DoD source databases (JPAS and CCMS) in order to investigate the extent to which the data agreed. Table 7 presents the results of this comparison. Since the data elements required for the discrepancy analysis were not returned in the Secondary Verification Response, Table 7 data were drawn from the 8,617 subjects verified in the Primary Response. The number of verified subjects with a data discrepancy is 623. It should be noted that 346 of these subjects reported their POB as the United States, accounting for most of the POB discrepancies (conditions 4 and 7).

Table 7
Discrepancies between Reported Data and USCIS Data (n = 8,617)

Condition	Discrepancy	n	%[13]
1	Last name differs	227	2.6
2	First name differs	19	.2
3	Date of birth differs	0	0
4	Place of birth differs	249	2.9
5	Name and date of birth differ	0	0
6	Name and place of birth differ	128	1.5
7	Date of birth & place of birth differ	0	0
8	Name & Date of birth & place of birth differ	0	0

Two modifications were required prior to the comparison: DoD and USCIS names were standardized to upper case and the USCIS 5-byte country code for POB was converted to the DoD 2-byte country code. Because of the complexity of cross-database matching, additional modifications should be made before names are compared, such as removing extraneous characters. To illustrate the overall complexity of matching, the types of discrepancies and concomitant examples[14] that were ruled out by manual review were as follows:

- Discrepant Names

  - Compound Names: Holmes vs. HolmesCraft; Olmos Ortgea vs. Olmos; SantosOrtega vs. OrtegaSantos

  - Spaces: VandeKamp vs. Van de Kamp

  - Dash: Smythe-Norton vs. Smythe Norton

  - Spelling: Gracia vs. Grazia

  - Middle Initial or Suffix in name field: Brown K vs. Brown; Smith Jr vs. Smith

  - Apostrophe: O'Halloran vs. OHalloran

  - Period: St. Sous vs. St Sous

  - Shortened: Sun vs. Sun Get

---

[13] Percentages may not sum to 100, due to rounding.
[14] Names are fictitious.

RESULTS

- Americanized: Joseph vs. Jose
  - Reported Alias matched USCIS
  - Reported Maiden name matched USCIS
- Discrepant birthdates: only the day portion of the date differed
- Discrepant Places of Birth
  - No POB reported by USCIS
  - Similar, but incorrect, country code in DoD data: NI for Niger, which is NG
  - Unknown POB in DoD Data
  - USCIS POB maps to Unknown

Overall, we were able to eliminate over 3,000 discrepancies through manual review. Many of the discrepancies can be eliminated by standardizing each name to eliminate spaces, apostrophes, and other extraneous characters prior to comparison. However, more sophisticated matching algorithms will be required to address compound names – particularly when the order of the name is reversed. Likewise, names which have been shortened can be addressed in matching, but only if the shortened name is a complete subset of the longer name. That is, Sun Get and Sun might match, but Cam and Cahguram might not. An immigrant population can exacerbate name-matching problems, as many adopt names that replace, rather than shorten or "Americanize," their given names.

Robust matching algorithms are imperative in systems where incoming data must be matched to an existing subject, either to avoid duplication of that subject or to attribute data to the appropriate subject. However, they should be applied judiciously in identity vetting, as minor modifications to identifying data may be employed to mask identity fraud. Thus, only modifications that remove extraneous characters are recommended. Clearly, all discrepancies will require review, as automated matching is literal and results in many false positives.

The match criteria used by USCIS were:

- ARN must match SAVE
- First character of the first name must match SAVE
- Month and year of DOB must match SAVE

Although these criteria obviate many of the automated name-matching problems, they place a greater burden on the end user to verify that the reported name, date of birth, and POB match SAVE. One of the advantages of the matching criteria, however, is that investigators and adjudicators are provided with an extra resource for identifying unreported aliases.

USCIS VERIFICATION AND DATA DISCREPANCIES BY SAMPLE CHARACTERISTIC

Verification results were inspected by each of the sample characteristics to determine whether verification varied by characteristic; Table 8 provides the tabulation of the USCIS Verification by sample characteristic. Because data discrepancies can only be determined for subjects verified in the Primary Response, both the number and percentage of subjects with discrepancies are provided in the table.

Only 4% of subjects were unverified; thus it is difficult to identify any relationship between sample characteristics and verification results. Within each characteristic, over 90% of the subjects were verified. Likewise, the percentage of discrepant data within each characteristic was small. The percentage of subjects with data discrepancies was greater than 10% for only two groups: Females (12.9%) and Dual Citizens (16.7%). Of the 302 women with data discrepancies, 220 occurred in the last name and were likely the result of name changes due to changes in marital status. The high percentage of discrepancies for dual citizens was based on too few subjects (one out of six) to draw conclusions.

DATA QUALITY ISSUES

Throughout the report, data quality or lack thereof has been an issue. Data quality refers to both data that are in error and data that should be present but are not. Specifically, the following data quality problems were encountered:

- ARNs that, once cleaned of extraneous characters, were blank
- ARNs that appeared to be made up because they were extremely low, sequential or repeating.
- ARNs that included free text
- ARN reported yet country of birth reported as the United States
- Missing dates of birth
- A large percentage of unknowns in critical fields such as Citizenship (11.1% of sample), Job Status (47.7% of sample population; but was improved by inferring Military from specific Investigation Types), and Eligibility (70% of unverified subjects)

RESULTS

<div align="center">

Table 8
USCIS Verification and Data Discrepancy by Sample Characteristic (n = 9983)

</div>

Sample Characteristic		Total n	Verified n	Verified %	Data Discrepancies n	Data Discrepancies %
Gender	Female	2341	2243	95.8	302	12.9
	Male	7629	7311	95.8	320	4.2
	Unknown	13	12	92.3	1	7.7
Age	< 19	3	3	100.0	0	.0
	19 - 21	230	216	93.9	14	6.1
	22 - 24	1926	1840	95.5	123	6.4
	25 - 29	5028	4834	96.1	330	6.6
	30 - 34	1381	1330	96.3	78	5.6
	35 - 39	731	699	95.6	39	5.3
	40 - 44	351	337	95.7	19	5.4
	45 - 50	203	186	91.6	16	7.9
	> 50	130	121	93.1	4	3.1
Investigation Type	NACLC	2606	2484	95.3	233	8.9
	NACLC-PR	33	31	93.9	2	6.1
	SSBI	141	133	94.3	3	2.1
	SSBI-PR	19	15	79.0	0	.0
	Other	6930	6665	96.2	366	5.3
	Unknown	254	238	93.7	19	7.5
Job Status	Civilian	132	128	97.0	9	6.8
	Contractor	249	231	92.8	16	6.4
	Military	4839	4638	95.8	324	6.7
	Unknown	4763	4569	95.8	274	5.8
Citizenship	U.S. Citizen	3909	3743	95.8	286	7.3
	Not a U.S. Citizen	4963	4766	96.0	266	5.4
	Dual Citizen	6	6	100.0	1	16.7
	Unknown	1105	1049	94.9	70	6.3

While the root cause of these issues may lie with flawed data entry – either on the part of the applicant or another performing the data entry – there are at least two areas where data entry problems could be reduced. The first is to place more emphasis on the ARN for individuals reporting a POB outside of the United States and the second is to allow a feedback mechanism into JPAS that would allow the correction of errant data.

A Procedural Guidance Message,[15] addressing the compliance with DoD guidelines on the quality and scope of citizenship data, was distributed to all Operations and Military Entrance Processing (MEP) Personnel in 2004. Effective June 1, 2004, it addressed the use of USCIS match results in the shipment of new recruits. Specifically, applicants born outside the United States must provide an ARN and that ARN and concomitant identifying data must match data in USCIS before processing can be completed. Therefore, subjects who entered the military after May 31, 2004, were required to provide a valid ARN and for that ARN, name, DOB, gender, and POB to match USCIS. This diligence should be applied to the personnel security process, particularly for nonmilitary applicants; military applicants should be required to provide a valid ARN on security questionnaires.

Much of the identifying data for civilians and the military are populated in JPAS via DoD Personnel Systems, with contractor identifying data populated from personnel security questionnaires. JPAS has required that errant data be corrected in these sources, which would then flow to JPAS. However, it is not clear that there is a feedback mechanism for errant data discovered during investigations to the appropriate personnel system: in most cases, the individual must contact his or her personnel office to implement changes. This leaves critical data in error with the corrective data concealed in investigative comments/summaries – when corrections are provided. Not only does this hamper research on personnel involved in the DoD Personnel Security Program, but it could impact continuing evaluation programs that rely on JPAS data.

---

[15] Procedural Guidance Message from HQ AFRS/RSO to All Operations/MEPS Personnel regarding Applicant Personal Data Verification. Signed by LtCol Daniel Woolever, USAF, Deputy Chief Operations Division.

RECOMMENDATIONS

IMPROVE THE VERIFICATION OF CITIZENSHIP AND ALIEN DATA

Current policy for vetting ARN information places full responsibility on front-line personnel to know when and how to request validation. It is at their discretion to rely solely on the information provided by applicants as recorded on documents that the applicants present vice corroborating the information through independent sources. If applicants have become naturalized since immigrating to the United States, personnel security policy and procedures do not even require recording of ARNs that were issued at time of immigration. Current policies and procedures reflect the importance placed by the personnel security system on documenting that subjects meet citizenship requirements but do little in the way of protecting against identity and application fraud on the part of those same subjects. In light of the availability independent corroboration of immigration and naturalization information through USCIS, better use of this information should be required.

- At a minimum, subjects born outside the United States to other than U.S. parents should be required to provide their ARN, even if they have subsequently become naturalized U.S. citizens.

- Use the SAVE system to validate ARN information for all personnel who have been issued ARNs, regardless of their citizenship status at the time of investigation.

- Where applicants have gained entry through granting of asylum, review their paperwork for requesting asylum.

- Adopt investigative guidelines for case expansion of investigations with unverified ARN information or verified ARN information with significant discrepancies. Ensure aliases that develop are documented and included in subsequent components of investigations.

- Where applicants are verified but with significant name discrepancies, determine why names were listed as they were and ensure developed aliases are included in other record checks.

- Because reporting the United States as the POB appears to be a relatively common data error, ensure that the true POB is obtained and correctly documented. Subjects reporting non-U.S. citizenship, dual citizenship, or foreign passports (current or expired) should be questioned as to their place of birth or citizenship status.

- Human Resource and Security officers should be trained and required to review documents, and review relevant portions of personnel security questionnaires and attest to their completeness and accuracy. As is done at the MEPS, background investigations should not proceed until ARNs are supplied.

Additionally, security questionnaires should include a question for aliens regarding the name in which their ARNs were issued. One would assume that the 'original' name would be reported in the Alias section of the form. However, multiple aliases

may exist, which would require multiple USCIS queries. An item specifically addressing the name associated with the ARN would allow investigators to check USCIS once, with the correct name. PERSEREC has made a similar recommendation for SSNs in the study addressing the verification of Social Security Numbers (SSN) via the Social Security Administration's Enumeration Verification System (Ainslie & Buck, in process). This approach has been adopted by the Internal Revenue Service (IRS) instructions for electronic filers. The IRS instructs filers to report their names *exactly* as they appear on their Social Security Cards, regardless of the names on their birth certificates.

DEVELOP AND IMPLEMENT INVESTIGATIVE PROCEDURES WHEN ARNS ARE NOT VERIFIED

If subjects cannot be verified, or they have been verified but significant discrepancies exist between SAVE and reported data, they should be interviewed to rule out errant data or to obtain the identifying data with which they entered the United States. If aliens cannot be verified via SAVE, then their investigations should not conclude until resolved.

Subject interviews to resolve ARN discrepancies should consist of reviewing applicants' Alien Registration Cards or Citizenship or Naturalization[16] Papers. Data on source documents should be compared to data that were sent for verification. If errors are with the data provided to USCIS, then they should be corrected and resubmitted.

If, on the other hand, applicants claim that the submitted data were correct, then they should be directed to the USCIS website where they can contact USCIS or find the nearest USCIS Field Office to resolve their cases.

Possible new investigative standards upon a finding of discrepancies between USCIS files and information provided by subjects are as follows:

- Request subjects to show their official Alien Registration Card to determine if discrepancies are due to data entry errors.

- If discrepancies cannot be resolved favorably for subjects, suspend investigations.

- Document reasons for suspension of investigations in JPAS, the SII, or other appropriate systems.

- Instruct subjects to resolve issues with USCIS.

- Resume processing only after subjects provide evidence of resolution and ARN can be verified.

---

[16] The ARN appears on the Naturalization Certificate, but is identified as either 'INS Registration Number' or 'CIS Registration Number.'

RECOMMENDATIONS

- If discrepancies cannot be resolved through USCIS or if subjects fail to follow through with efforts to resolve discrepancies, notify appropriate authorities[17] and document status in JPAS.

- Identify appropriate authorities for resolving potentially fraudulent ARNs.

For cases that cannot be resolved as data entry errors on the part of DoD or USCIS, DUSD(CI&S) will need to develop policy for investigation, adjudication, and flagging those subjects in the event they try to reapply at a later time.

DEVELOP AND IMPLEMENT INVESTIGATIVE PROCEDURES WHEN ARNS ARE VERIFIED WITH DISCREPANCIES

For a given ARN, SAVE checks require only matching the month and day of the date of birth and the first letter of the first name. Consequently, SAVE may uncover days of birth, unreported aliases, and countries of birth that differ from those reported by subjects. Therefore:

- Verified responses should be reviewed by investigators or adjudicators to ensure that data on file for immigrant subjects in DoD are consistent with those on record with USCIS.

- When discrepancies are deemed significant, they should trigger expanded investigation and special interviews with subjects at the same level as what would be required when subjects are reported as unverified by USCIS.

- Given the importance of identity vetting and the ability to verify an individual across widely varying systems, it is worth considering a modification to the manner in which aliases are stored. If an alias has been used to verify alien status, then that alias should be so marked. Likewise, if an alias (or maiden name) was used to verify an SSN, it should be so marked. This would preclude problems with later verification attempts and provide the documentation necessary to support quality personnel security investigations and adjudications.

IMPROVE DATA QUALITY

The Government Accountability Office (GAO-07-310, 2007) identified four actions that DoD must take to improve its personnel security clearance program. One of these actions was to implement procedures to eliminate documentation problems. Providing a mechanism to correct errant data would greatly improve electronic documentation and could result in fewer requests for supporting documentation.

Review Data Holdings

The growing pervasiveness of database sources that can be used for research and management decisionmaking means that more attention must be focused on data quality. Researchers must become intimately familiar with the data sources in

---

[17] Exactly who will be designated as appropriate authorities to deal with suspected identity fraud will need to be determined by the DUSD(CI&S) Director of Security.

order to correctly interpret anomalous results. Since the quality of documentation differs dramatically from one source to the next, it is imperative that users obtain data samples and review them in light of matching requirements and the provided documentation. Discrepancies should be discussed with the data holder, in order to reach an understanding of those discrepancies and deal with them appropriately. The importance of reviewing and analyzing the data cannot be overstated.

Implement Quality Controls

Data quality problems can be obviated with data entry controls. When entering data using an automated tool, users should be allowed to enter unconstrained data only where absolutely necessary (SSN, ARN, name, DOB, addresses, etc.). While the content of these fields cannot be constrained, the length of SSN, ARN, and DOB can, ensuring at least this degree of conformance. Additionally, critical data elements, such as SSN and ARN, should require double entry to obviate errors.

For coded values, users should be provided lists of allowed values from which to choose. And, when the relationship between two fields is unambiguous, data entry should be so constrained. For example, users who indicate they were not born in the United States should be guided to the entry of an ARN. Likewise, users who enter ARNs and indicate that they were born in the United States should be instructed to review the conflicting entries.

Until cross-referencing of related data elements can be automated security officers should fill the gap with careful review of related fields and increased emphasis on immigration data. Sponsoring security officers should implement a program to review all data entry and obtain ARNs before submitting paperwork for the non-U.S. citizens.

Future Directions

DoD should also develop policy and procedures for vetting ARN information provided by applicants against ARN information already on file in DoD systems. Doing so would defend against subjects borrowing each other's personal identifiers or using the same criminal sources that sell fraudulent identifications to willing buyers.

PERSEREC is also working with USCIS to determine capacities for low-cost, high-volume verification of naturalization numbers. These currently entail either manual review of documents or investigators personally contacting USCIS. For the same reasons that independent verification of ARN information is essential for determining security clearance eligibility and security risk as indicated by potential applicant fraud, independent and automated checks of all naturalization information would also be prudent.

Given the importance of identity vetting and the ability to verify an individual across widely-varying systems, it is worth considering a modification to the manner

RECOMMENDATIONS

in which aliases are stored. If an alias has been used to verify alien status, then that alias should be so marked. Likewise, if an alias (or maiden name) was used to verify an SSN, it should be so marked. This would preclude problems with later verification attempts and provide the documentation necessary to support quality personnel security investigations and adjudications.

Finally, while the USCIS Secondary Verification process was not designed to provide identifying information, modifying the response to include the same data that are returned in the Primary Verification would aid in identifying errant ARNs, data discrepancies and aliases, and support the documentation needs of the personnel security program.

REFERENCES

Ainslie, F.M., & Buck, K.R. (in process). *Social security number (SSN) verification via the Social Security Administration's Enumeration Verification System (EVS).*

Government Accountability Office. (2007). *High-risk series: An update* (GAO-07-310). Washington, DC: Author.

U.S. Department of Homeland Security. (2006). *Customer Processing System (CPS) Automated Secondary Batch Access Method Interface Control Document,* Agency Version, Draft v38 (ELCN). Washington, DC: Author.