Privacy Office
U.S. Department of Homeland Security
Washington, DC 20528
202-343-1717, pia@dhs.gov
www.dhs.gov/privacy

PRIVACY THRESHOLD ANALYSIS (PTA)

This form is used to determine whether
a Privacy Impact Assessment is required.

Please use the attached form to determine whether a Privacy Impact Assessment (PIA) is required under the E-Government Act of 2002 and the Homeland Security Act of 2002.

Please complete this form and send it to your component Privacy Office. If you do not have a component Privacy Office, please send the PTA to the DHS Privacy Office:

Senior Director, Privacy Compliance
The Privacy Office
U.S. Department of Homeland Security
Washington, DC 20528
Tel: 202-343-1717

PIA@hq.dhs.gov

Upon receipt from your component Privacy Office, the DHS Privacy Office will review this form. If a PIA is required, the DHS Privacy Office will send you a copy of the Official Privacy Impact Assessment Guide and accompanying Template to complete and return.

A copy of the Guide and Template is available on the DHS Privacy Office website, www.dhs.gov/privacy, on DHSConnect and directly from the DHS Privacy Office via email: pia@hq.dhs.gov, phone: 202-343-1717.

PRIVACY THRESHOLD ANALYSIS (PTA)

SUMMARY INFORMATION

Project or Program Name:	**Giant Oak Search Technology (GOST)**		
Component:	Immigration and Customs Enforcement (ICE)	**Office or Program:**	Homeland Security Investigations (HSI), National Security Investigations Division (NSID), and International Operations
Xacta FISMA Name (if applicable):	**N/A**	**Xacta FISMA Number (if applicable):**	**N/A**
Type of Project or Program:	IT System	**Project or program status:**	Existing
Date first developed:	N/A	**Pilot launch date:**	September 4, 2014
Date of last PTA update	**N/A**	**Pilot end date:**	Click here to enter a date.
ATO Status (if applicable)	In progress	**ATO expiration date (if applicable):**	**TBD**

PROJECT OR PROGRAM MANAGER

Name:	(b)(6); (b)(7)(C)		
Office:	CTCEU	**Title:**	National Program Manager
Phone:	(703) 23(b)(6); (b)(7)(C)	**Email:**	(b)(6); (b)(7)(C) @ice.dhs.gov

INFORMATION SYSTEM SECURITY OFFICER (ISSO) (IF APPLICABLE)

Name:	(b)(6); (b)(7)(C)		
Phone:	(703) 842-(b)(6); (b)(7)(C)	**Email:**	(b)(6); (b)(7)(C) @giantoak.com

SPECIFIC PTA QUESTIONS

1. Reason for submitting the PTA: New PTA

HSI NSID and International Operations identify and eliminate vulnerabilities in the nation's security through programs including the Counterterrorism and Criminal Exploitation Unit (CTCEU) and the Visa Security Program (VSP). The Office of Professional Responsibility (OPR) ensures the integrity and safety of the ICE workforce. CTCEU, OPR, and VSP utilize commercial vendor Giant Oak's off-the-shelf, web-based platform called Giant Oak Search Technology (GOST) to monitor publicly available information via open-source websites and publicly available social media accounts for investigative leads. Targets of investigation are non-US persons, but information regarding associates who may be US citizens or LPRs may be collected in the course of an investigation.

HSI Use

CTCEU and VSP users (hereafter, HSI users) access GOST through the Giant Oak web portal. HSI users upload data schemas to the web portal as a CSV file containing 30-50 data points about persons of interest from HSI cases in Leadtrac. Available data fields include biographical information, contact information, college and course of study information

Using the data points included in the HSI data schemas, GOST searches the internet for publicly available information from open source and social media sites belonging to the subjects. GOST searches three types of datasets:

1. Open Source information: All information that would be searchable using public search engines.
2. Proprietary/Deep Web information: Information that at one time was publicly available, but was harvested by GOST and stored due to its determined importance to GOST searches and the likelihood that it would be de-indexed and therefore unsearchable in the future.
3. Paid data sources: Publicly available information collated by third party search engines, such as LexisNexis, ThomsonReuters, Transunion, and international news organizations. The ultimate list of subscriptions is proprietary to GOST and unknown to ICE. The sources of the information from those returns from third party search engines will be noted in GOST.

Social Media accounts and postings could appear in a search of any of the three types of data. Social media accounts retrieved include, but are not limited to: Facebook, Google+, LinkedIn, Pinterest, Tumblr, Instagram, VK, Flickr, Myspace, and Twitter. GOST does not use any social media account to access information. It also does not violate any site paywall restrictions or circumvent any social media account's privacy settings.

GOST returns its findings to the web portal, broken into five categories:

(b)(7)(E)

(b)(7)(E)

Once GOST has returned its search results to the Giant Oak web portal, HSI users review the information to determine whether the information returned pertains to the correct individual, and whether the information is of value. Users can also run reverse image searches of photographs found by GOST to determine if the same image is posted elsewhere on the internet. GOST also allows for a user to run a facial recognition algorithm on all images collected in a search for an individual to help determine if the individual is depicted in associated images. The facial recognition algorithm can only be conducted on images already returned by GOST for a particular subject.

Users will select relevant information in the results that will then automatically be added to a report that can be exported from the web portal. The act of selecting information by a user as valuable informs the GOST search algorithms for Ranking future searches.

CTCEU uses Giant Oak to assist the visa overstay/lead vetting process, primarily to discover location information of a subject, but also for evidence of criminal activity or national security concerns. CTCEU also monitors student visa holders to determine whether students have changed their field of study to a sensitive area that may have military applications or counterproliferation concerns, as defined by the Department of State. GOST is used as a monitoring tool, and CTCEU's open source team (OST) will manually search open source materials to exhaust leads prior to entering an individual into GOST. When GOST alerts CTCEU of new open source material that may be posted, OST will re-initiate checks against government systems and again manually search for additional open source information prior to entering information into LeadTrac. Subject information downloaded or documented in LeadTrac will additionally be stored in a secure Microsoft Access database located on a DHS shared drive for use in assigning tasks to OST analysts.

(b)(7) users parse out a subset of visa application data and upload this data to the Giant Oak web portal mentioned above. All information determined to be of value will be documented manually in (b)(7)(E) - (b)(7)(E) can also refer to CTCEU for continuous monitoring of a subset of subjects (i.e., F, M, J, B1, and B2 visa applicants from pre-designated consular posts). CTCEU will use Giant Oak to engage in continuous evaluation of the subject. The vetting is confined to searches for evidence of criminal and national security concerns. CTCEU users will not actively vet these subjects, but will set threshold alerts within GOST.

Excluding ad hoc search results, which are removed on a weekly basis, information will be removed from the Giant Oak web portal when HSI determines that the individual to whom it pertains is no longer subject to analysis

HSI users may also export the search results, excluding photographs, to a CSV file. This CSV file cannot be downloaded, but may be burned onto a CD and then transferred to a LeadTrac-attributed computer. Photographs may separately be saved to a CD, and then transferred to a LeadTrac-attributed computer.

OPR Use

OPR uses GOST for purposes of situational awareness and protection of ICE property and personnel. Instead of searching for individuals through GOST, OPR sets search parameters based on keyword searches. OPR has determined a set of words that may imply violence toward ICE as well as names of ICE leaders or personnel determined to be under threat. GOST issues daily reports of posts that contain the pre-determined keywords. The daily reports are not saved within GOST, and OPR personnel do not have user accounts for GOST. Reports only contain a screenshot of the post that contained the keyword, to include the account name of the poster. That report is analyzed by OPR personnel and forwarded on to OPR's investigative unit if it is determined that the threat is credible. OPR will store the credible reports on a shared drive. If any posts deemed to be credible lead to an investigation by OPR, they will upload the report to their case management system, U.S. Customs and Border Protection's (CBP's) Joint Integrity Case Management System (JICMS).

2. Does this system employ any of the following technologies: *If you are using any of these technologies and want coverage under the respective PIA for that technology please stop here and contact the DHS Privacy Office for further guidance.*	□ Closed Circuit Television (CCTV) ☒ Social Media □ Web portal[1] (e.g., SharePoint) □ Contact Lists □ None of these

3. From whom does the Project or Program collect, maintain, use, or disseminate information?	□ This program does not collect any personally identifiable information[2]

[1] Informational and collaboration-based portals in operation at DHS and its components that collect, use, maintain, and share limited personally identifiable information (PII) about individuals who are "members" of the portal or "potential members" who seek to gain access to the portal.

[2] DHS defines personal information as "Personally Identifiable Information" or PII, which is any information that permits the identity of an individual to be directly or indirectly inferred, including any information that is linked or linkable to that individual, regardless of whether the individual is a U.S. citizen, lawful permanent resident, visitor to the U.S., or employee or contractor to the Department. "Sensitive PII" is PII, which if lost, compromised, or disclosed without authorization, could result in substantial harm, embarrassment, inconvenience, or unfairness to an individual. For the purposes of this PTA, SPII and PII are treated the same.

Please check all that apply.	☒ Members of the public ☒ DHS employees/contractors (list components): ICE ☒ Contractors working on behalf of DHS ☐ Employees of other federal agencies

4. What specific information about individuals is collected, generated or retained?

Members of the Public:

Information contained in the HSI data schemas that are uploaded to the Giant Oak web portal includes: name; date of birth; country of birth and country of citizenship; aliases information; affiliate and associate names, dates of birth, relation, countries, and occupations; vehicle and license plate information; driver's license or state identification number; LeadTrac number; social security number (SSN); address information; phone number; e-mail address; IP address; web identity (i.e. account name, social media handle); and college and course of study information. GOST's search results depend on the amount of information available on open-source and social media for each subject, but all results can be broken into five categories:

1. Web locations at which information about the person of interest was found, including a short summary of the finding, an image of the site, and a link to the site;
2. Links to open-source and social media accounts belonging to the person of interest;
3. Images of the person of interest;
4. Social Graph: a visualization of connections of family and known associates derived from a GOST search; and
5. Location information, including addresses associated with the person of interest.

ICE Employees/Contractors

GOST collects username and contact information from GOST users to allow for access and authorization to use the system.

4(a) Does the project, program, or system retrieve information by personal identifier?	☐ No. Please continue to next question. ☒ Yes. If yes, please list all personal identifiers used: Data stored in GOST is retrievable by Name, LeadTrac ID number, immigration status, or assignee (ICE employee or contractor).
4(b) Does the project, program, or system use Social Security Numbers (SSN)?	☐ No. ☒ Yes.

4(c) If yes, please provide the specific legal basis and purpose for the collection of SSNs:	Immigration and Nationality Act sections 328 and 329; U.S.A. PATRIOT Act of 2001, Public Law 107–56; the Border Security Act, Public Law 107–173; and the Implementing Recommendations of the 9/11 Commission Act of 2007, Public Law 110-53
4(d) If yes, please describe the uses of the SSNs within the project, program, or system:	CTCEU uses SSNs to ensure clarity in identity resolution of the subjects of investigation and their associates. The more certain CTCEU can be as to the actions and connections of the subjects of investigation, the more likely the investigation will result in a viable lead. CTCEU's primary focus is locating and tracking subjects of investigations, which is directly affected by the quality of identity resolution.
4(e) If this project, program, or system is an information technology/system, does it relate solely to infrastructure? *For example, is the system a Local Area Network (LAN) or Wide Area Network (WAN)?*	☒ No. Please continue to next question. ☐ Yes. If a log kept of communication traffic, please answer the following question.

4(f) If header or payload data[3] is stored in the communication traffic log, please detail the data elements stored.
N/A

5. Does this project, program, or system connect, receive, or share PII with any other DHS programs or systems[4]?	☐ No. ☒ Yes. If yes, please list: GOST does not have a system-to-system connection with any system. Data elements used for GOST searches are derived from LeadTrac and (b)(7)(E) -(b)(7)(E) Investigative information obtained

[3] When data is sent over the Internet, each unit transmitted includes both header information and the actual data being sent. The header identifies the source and destination of the packet, while the actual data is referred to as the payload. Because header information, or overhead data, is only used in the transmission process, it is stripped from the packet when it reaches its destination. Therefore, the payload is the only data received by the destination system.

[4] PII may be shared, received, or connected to other DHS systems directly, automatically, or by manual processes. Often, these systems are listed as "interconnected systems" in Xacta.

	from GOST searches is manually uploaded to LeadTrac, JICMS, and (b)(7)(E)
6. Does this project, program, or system connect, receive, or share PII with any external (non-DHS) partners or systems?	☒ No. ☐ Yes. If yes, please list:
6(a) Is this external sharing pursuant to new or existing information sharing access agreement (MOU, MOA, LOI, etc.)?	Please describe applicable information sharing governance in place: N/A.
7. Does the project, program, or system provide role-based training for personnel who have access in addition to annual privacy training required of all DHS personnel?	☒ No. ☐ Yes. If yes, please list:
8. Per NIST SP 800-53 Rev. 4, Appendix J, does the project, program, or system maintain an accounting of disclosures of PII to individuals who have requested access to their PII?	☒ No. What steps will be taken to develop and maintain the accounting: Information generated through GOST will be manually uploaded to LeadTrac and (b)(7)(E) - (b)(7)(E) and any disclosures would occur from LeadTrac/(b)(7)(E) rather than GOST. ☐ Yes. In what format is the accounting maintained:
9. Is there a FIPS 199 determination?[4]	☐ Unknown. ☐ No. ☒ Yes. Please indicate the determinations for each of the following: Confidentiality: ☐ Low ☒ Moderate ☐ High ☐ Undefined Integrity: ☐ Low ☒ Moderate ☐ High ☐ Undefined Availability:

[4] FIPS 199 is the Federal Information Processing Standard Publication 199, Standards for Security Categorization of Federal Information and Information Systems and is used to establish security categories of information systems.

	☐ Low ☒ Moderate ☐ High ☐ Undefined

PRIVACY THRESHOLD REVIEW

(TO BE COMPLETED BY COMPONENT PRIVACY OFFICE)

Component Privacy Office Reviewer:	(b)(6); (b)(7)(C)
Date submitted to Component Privacy Office:	**February 19, 2019**
Date submitted to DHS Privacy Office:	May 8, 2019

Component Privacy Office Recommendation:
Please include recommendation below, including what new privacy compliance documentation is needed.

GOST is a privacy sensitive system in that it collects, uses, and maintains PII from members of the public. All searches executed by GOST are in compliance with the ICE Criminal and Administrative Immigration Law Enforcement SMOUT executed 12/5/2016. ICE recommends updating this SMOUT to reflect the addition of the LeadTrac PIA and SORN in its operational use, as well as clarifying OPR's routine use of social media for officer safety.

The forthcoming ICE Social Media PIA will cover the privacy risks inherent in using GOST. Interim coverage is provided by:

- DHS/ICE/PIA-044 LeadTrac PIA, which assesses the privacy risks of collecting open source data for investigative leads
- DHS/CBP/PIA-044 Joint Integrity Case Management System, which discusses the inclusion of open source data in the system.
- DHS/ICE/PIA-011(a) Visa Security Program Tracking System -Network Version 2.0 covers the system collecting open source information in VSP's vetting process.

The information input into GOST comes from HSI's investigative case files. Results derived from GOST are evidentiary in nature and used in ICE investigations. ICE recommends

- Current coverage under DHS/ICE-009 – External Investigations SORN, as all data collected and maintained is for the purpose of supporting investigations of criminal activity. ICE will include in its next update that open source information and social media posts as categories of records for further transparency.
- Coverage for information entered into Leadtrac under DHS/ICE-015 LeadTrac SORN

- Coverage for information entered into (b)(7)(E) under (b)(7)(E) (b)(7)(E) SORN, which covers public records obtained during a visa security review.

(TO BE COMPLETED BY THE DHS PRIVACY OFFICE)

DHS Privacy Office Reviewer:	(b)(6); (b)(7)(C)
PCTS Workflow Number:	**1180593**
Date approved by DHS Privacy Office:	June 12, 2019
PTA Expiration Date	June 12, 2022

DESIGNATION

Privacy Sensitive System:	Yes If "no" PTA adjudication is complete.
Category of System:	IT System If "other" is selected, please describe: Click here to enter text.
Determination:	☐ PTA sufficient at this time. ☐ Privacy compliance documentation determination in progress. ☐ New information sharing arrangement is required. ☐ DHS Policy for Computer-Readable Extracts Containing Sensitive PII applies. ☐ Privacy Act Statement required. ☒ Privacy Impact Assessment (PIA) required. ☒ System of Records Notice (SORN) required. ☐ Paperwork Reduction Act (PRA) Clearance may be required. Contact your component PRA Officer. ☐ A Records Schedule may be required. Contact your component Records Officer.
PIA:	**System covered by existing PIA** If covered by existing PIA, please list: (b)(7)(E) Forthcoming ICE Social Media PIA; DHS/ICE/PIA-044 LeadTrac and forthcoming appendix updates; DHS/ICE/PIA-011(a) Visa Security Program Tracking System -Network Version 2.0; DHS/CBP/PIA-044 JICMS

SORN:	SORN update is required. If covered by existing SORN, please list: Forthcoming update to DHS/ICE-009 External Investigations January 5, 2010 75 FR 404; DHS/ICE-012 (b)(7)(E) September 30, 2009 74 FR 50228; DHS/ICE-015 LeadTrac System of Records, August 9, 2016, 81 FR 52700

DHS Privacy Office Comments:
Please describe rationale for privacy compliance determination above.

ICE is submitting this PTA to discuss the Counterterrorism and Criminal Exploitation Unit (CTCEU), (b)(7)(E) and Office of Professional Responsibility (OPR) use of Giant Oak Search Technology (GOST). CTCEU and (b)(7) GOST to monitor publicly available information via open-source websites and publicly available social media accounts for investigative leads. Targets of investigation are non-U.S. persons, but information regarding associates who may be U.S. citizens or LPRs may be collected in the course of an investigation.

OPR uses GOST for purposes of situational awareness and protection of ICE property and personnel. Instead of searching for individuals through GOST, OPR sets search parameters based on keyword searches. GOST issues daily reports of posts that contain the pre-determined keywords. The daily reports are not saved within GOST, and OPR personnel do not have user accounts for GOST.

The DHS Privacy Office finds this is a privacy sensitive system, requiring PIA coverage. ICE is currently drafting a Social Media PIA to cover ICE uses of open source and social media information, and that PIA will include discussion of GOST and should also include specific discussion of the use of facial recognition in searching. Coverage will also be provided by the forthcoming PATRIOT PIA, and forthcoming updates to the appendices of DHS/ICE/PIA-044 LeadTrac.

SORN coverage is also required, as information is retrieved by identifier. ICE is currently updating DHS/ICE-009 External Investigations, and should include social media and open source information. Coverage for information entered into LeadTrac is provided by DHS/ICE-015 LeadTrac, and coverage for information entered into (b)(7)(E) is provided by DHS/ICE-012 (b)(7)(E)

ICE should submit an updated the ICE Criminal and Administrative Immigration Law Enforcement SMOUT, and PRIV recommends submission of a separate SMOUT for OPR's situational awareness use of social media.

This PTA will expire in one year. Two additional years of coverage will be provided upon completion of the required PIA and SORN updates.

After a review of recently published documents, PRIV finds that coverage is provided by the recently published (b)(7)(E) DHS/ICE-009 External Investigations is in the process of being updated to include social media and is near completion, and ICE has recently submitted appendix updates to the LeadTrac PIA.

More comprehensive coverage for ICE use of social media will continue to be provided by the forthcoming ICE social media PIA. **The DHS/ICE/PIA-054 ICE Use of Facial Recognition Services was recently published which outlines the use of facial recognition services (FRS) that require the collection, maintenance, and use of PII.**