PRIVACY THRESHOLD ANALYSIS (PTA)

This form will be used to determine whether a Privacy Impact Assessment (PIA), System of Records Notice (SORN), or other privacy compliance documentation is required under the E-Government Act of 2002, the Homeland Security Act of 2002, the Privacy Act of 1974, or DHS policy.

Please complete this form and send it to your Component Privacy Office. If you are unsure of your Component Privacy Office contact information, please visit https://www.dhs.gov/privacy-office-contacts. If you do not have a Component Privacy Office, please send the PTA to the DHS Privacy Office:

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

PIA@hq.dhs.gov

Your Component Privacy Office will submit the PTA on behalf of your office. Upon receipt from your Component Privacy Office, the DHS Privacy Office will review this form. If a PIA, SORN, or other privacy compliance documentation is required, your Component Privacy Office, in consultation with the DHS Privacy Office, will send you a copy of the template to complete and return.

For more information about the DHS Privacy compliance process, please see https://www.dhs.gov/compliance. A copy of the template is available on DHS Connect at (b)(7)(E) or directly from the DHS Privacy Office (b) (6)

PRIVACY THRESHOLD ANALYSIS (PTA)

SUMMARY INFORMATION

Project, Program, or System Name:	Matroid Data Annotation		
Component or Office:	Customs and Border Protection (CBP)	**Office or Program:**	Office of the Commissioner – Artificial Intelligence (AI) Center of Innovation (COI)
FISMA Name (if applicable):	N/A	**FISMA Number (if applicable):**	N/A
Type of Project or Program:	New project	**Project or program status:**	Pilot
Date first developed:	February 20, 2024	**Pilot launch date:**	February 20, 2024
Date of last PTA update	N/A	**Pilot end date:**	February 20, 2025
ATO Status (if applicable):[1]	Not started	**Expected ATO/ATP/OA date (if applicable):**	Click here to enter a date.

PROJECT, PROGRAM, OR SYSTEM MANAGER

Name:	(b) (6) (b) (7) (c)		
Office:	Office of the Assistant Commissioner	**Title:**	Lead ITSPEC
Phone:	(b) (6) (b) (7) (c)	**Email:**	(b) (6) (b) (7) (c)@cbp.dhs.gov

INFORMATION SYSTEM SECURITY OFFICER (ISSO) (IF APPLICABLE)

Name:	(b) (6) (b) (7) (c)		
Phone:	(b) (6) (b) (7) (c)	**Email:**	(b) (6) (b) (7) (c)@associates.cbp.dhs.gov

[1] The DHS OCIO has implemented a streamlined approach to authorizing an Authority to Operate (ATO), allowing for rapid deployment of new IT systems and initiate using the latest technologies as quickly as possible. This approach is used for selected information systems that meet the required eligibility criteria in order to be operational and connect to the network. For more information, see (b)(7)(E)

SPECIFIC PTA QUESTIONS

1. Reason for submitting the PTA: New PTA

Customs and Border Protection (CBP) is submitting this new PTA for the CBP Innovation Team, to test the use of Matroid Software. (b) (4) [redacted] The Artificial Intelligence (AI) Center of Innovation (COI) Office of Commissioner will work with Matroid Inc., the developer of the Matroid software, with the intent of further developing the Matroid software to meet the CBP operational requirements.

Purpose

The Matroid software is a proof of concept (POC) that will enable CBP to label and annotate data while creating and sharing detectors, which are (b) (4) [redacted] Matroid does not create its own video, users either upload previously recorded images or video into the Matroid application or users directly link existing camera feeds into the application. CBP agents and officers may view detection information via variety of reports and alert notifications to process and identify objects and people. Detection data is also available through Matroid's developer API and language-specific clients, so CBP applications can be integrated with the computer vision provided.

The operational need for this capability stems from the incredible volume of imagery collected by CBP across its mission space every day. In almost all cases, various imagery streams must be monitored and interrogated by CBP employees to identify anomalies, threats, and other mission critical information. Monitoring imagery for specific objects is a highly fatiguing task, and in most cases is better suited for computers than humans.

CBP Office of Field Operations (OFO) will provide 440 images from the (b) (7)(E) [redacted] and 1,001,000 images from the (b) (7)(E) [redacted]. Only CBP will have access to the data within the Matroid software application. The image files will be fully sanitized, with metadata and PII salt hashed to limit the possibility of re-identification. Both the (b) (7)(E) [redacted] data will be retained within the Matroid software application until the proof of concept is completed. Once the proof of concepts is completed, data will be deleted through a manual process. CBP data owners will be informed following deletion of the data within the Matroid software.

Matroid software will run on dedicated CBP owned and managed (b) (4) [redacted] To evaluate the development of the Matroid software, the CBP AI COI in conjunction with CBP agents and officers will pilot the Matroid software across multiple CBP offices to effectively determine the impact of its capabilities on CBP operations. Only select authorized CBP users will have access to any data input to and output by Matroid. (b) (4) [redacted] The Matroid software is permitted on the CBP Technical Reference Model (TRM) portal.

As of right now, the Matroid containers are using a (b) (7)(E) image for configuring and training of the detectors to allow for an expedited timeline, which does not currently have the Nessus agent installed. However, the server that Matroid is installed on has the Nessus agent installed.

Authorized CBP users will log into the Matroid software and upload select CBP data to be used in these data annotation activities and establish labeled data sets for use by CBP's Artificial Intelligence (AI) applications. (b) (4)

How it Works:

1. (b) (4)

Once the POC is completed, data will be deleted through a manual process and conformation will be provided to the Information System Security Officer (ISSO) and Security Technology Policy (STP). CBP data owners will be informed following deletion of the data within the Matroid Software.

AI COI will use this PTA to conduct Matroid software POC investigation and testing for data labeling and annotation Services, Machine Learning (ML), Model training, and Model testing and evaluation in support of the CTO. Once the AICOI has successfully tested the full capability of Matroid software data labeling and annotation capability, AI COI will seek an ATO.

2. From whom does the Project, Program, or System collect, maintain, use, or disseminate information?	☐ This project does not collect, collect, maintain, use, or disseminate any personally identifiable information[2]

[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

Please check all that apply.	☒ Members of the public ☒ U.S. Persons (U.S citizens or lawful permanent residents) ☒ Non-U.S. Persons ☒ DHS Employees/Contractors (list Components): **CBP** ☐ Other federal employees or contractors (list agencies): *Click here to enter text.*
2(a) Is information meant to be collected from or about sensitive/protected populations?	☒ No ☐ 8 USC § 1367 protected individuals (e.g., T, U, VAWA)[3] ☐ Refugees/Asylees ☐ Other. Please list: *Click here to enter text.*

3. What specific information about individuals is collected, maintained, used, or disseminated?

CBP collects imagery from a wide spectrum of sources each of which is documented in existing privacy documentation and appropriate systems documentation (ex. PTAs, PIAs, and SORNs). Examples include, but are not limited to:

– (b) (7)(E)

These photos will be uploaded into Matroid. The image files will be fully sanitized, with metadata and PII salt hashed to limit the possibility of re-identification.

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.

[3] This involves the following types of individuals: T nonimmigrant status (Victims of Human Trafficking), U nonimmigrant status (Victims of Criminal Activity), or Violence Against Women Act (VAWA). For more information about 1367 populations, please see: DHS Management Directive 002-02, Implementation of Section 1367 Information Provisions, *available at* (b)(7)(E)

Matroid does not create its own still imagery or video. Instead, users create "detectors" in the Matroid application and existing video is run through the software to make detections on behalf of the user.

Users can create ML Models or customize models in the Matroid software and existing video is ran through the software to make detections on behalf of the user.

- User Email (CBP email account)
- User Password

Additionally, all user interactions are logged.

3(a) Does this Project, Program, or System collect, maintain, use, or disseminate Social Security numbers (SSN) or other types of stand-alone sensitive information?[4] If applicable, check all that apply.

☐ Social Security number

☐ Alien Number (A-Number)

☐ Tax Identification Number

☐ Visa Number

☐ Passport Number

☐ Bank Account, Credit Card, or other financial account number

☐ Social Media Handle/ID

☐ Driver's License/State ID Number

☒ Biometric identifiers *(e.g., FIN, EID)*

☐ Biometrics.[5] *Please list modalities (e.g., fingerprints, DNA, iris scans): Click here to enter text.*

☐ Other. *Please list: Click here to enter text.*

3(b) Please provide the specific legal basis for the collection of SSN:	N/A

3(c) If the SSN is needed to carry out the functions and/or fulfill requirements of the Project, System, or Program, please explain why it is necessary and how it will be used.

N/A

3(d) If the Project, Program, or System requires the use of SSN, what actions are being taken to abide by Privacy Policy Instruction 047-01-010, *SSN Collection and Use Reduction*,[6] which requires the use of privacy-enhancing SSN alternatives when there are technological, legal, or regulatory limitations to eliminating the SSN? Note: *even if you are properly authorized to collect*

[4] Sensitive PII (or sensitive information) is PII that if lost, compromised, or disclosed without authorization, could result in substantial harm, embarrassment, inconvenience, or unfairness to an individual. More information can be found in the DHS Handbook for Safeguarding Sensitive Personally Identifiable Information, *available at* https://www.dhs.gov/publication/handbook-safeguarding-sensitive-personally-identifiable-information.

[5] If related to IDENT/HART and applicable, please complete all Data Access Request Analysis (DARA) requirements. This form provides privacy analysis for DHS' IDENT, soon to be HART. The form replaces a PTA where IDENT is a service provider for component records. PRIV uses this form to better understand how data is currently shared, will be shared and how data protection within IDENT will be accomplished. IDENT is a biometrics service provider and any component or agency submitting data to IDENT is a data provider.

[6] *See* https://www.dhs.gov/publication/privacy-policy-instruction-047-01-010-ssn-collection-and-use-reduction.

SSNs, you are required to use an alternate unique identifier. If there are technological, legal, or regulatory limitations to eliminating the SSN, privacy-enhancing alternatives should be taken, such as masking, truncating, or encrypting the SSN, or blocking the display of SSNs in hard copy or digital formats.
N/A

4. How does the Project, Program, or System retrieve information?	☐ By a unique identifier.[7] Please list all unique identifiers used: *Click here to enter text.* ☒ By a non-unique identifier or other means. Please describe: e.g. still images and videos primarily by filename (i.e. object key) but can also be retrieved based on date created and time

5. What is the records retention schedule(s) for the information collected for each category type (include the records schedule number)? *If no schedule has been approved, please provide proposed schedule or plans to determine it.* *Note: If no records schedule is in place or are unsure of the applicable records schedule, please reach out to the appropriate Records Management Office.*[8]	GRS 3.1, item 11.5 years after the CBP system or record is no longer in use
5(a) How does the Project, Program, or System ensure that records are disposed of or deleted in accordance with the retention schedule (e.g., technical/automatic purge, manual audit)?	Manual Audit

6. Does this Project, Program, or System connect, receive, or share PII with any other DHS/Component projects, programs, or systems?[9]	☐ No. ☒ Yes. If yes, please list:

[7] Generally, a unique identifier is considered any type of "personally identifiable information," meaning any information that permits the identity of an individual to be directly or indirectly inferred, including any other information which 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.

[8] *See* (b)(7)(E)

[9] 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 IACS.

	Images will be collecting data to use with Matroid that may contain facial, license plates, or other identifiable imagery from CBP.
7. Does this Project, Program, or System connect, receive, or share PII with any external (non-DHS) government or non-government partners or systems?	☒ No. ☐ Yes. If yes, please list: *Click here to enter text.*
8. Is this sharing pursuant to new or existing information sharing agreement (MOU, MOA, LOI, RTA, etc.)? *If applicable, please provide agreement as an attachment.*	N/A Please describe applicable information sharing governance in place: N/A
9. Does the Project, Program, or System or have a mechanism to track external disclosures of an individual's PII?	☒ No. What steps will be taken to develop and maintain the accounting: *Click here to enter text.* ☐ Yes. In what format is the accounting maintained: *Click here to enter text.*

10. Does this Project, Program, or System use or collect data involving or from any of the following technologies:	☐ Social Media ☒ Advanced analytics[10] ☐ Live PII data for testing ☐ No

[10] The autonomous or semi-autonomous examination of Personally Identifiable Information using sophisticated techniques and tools to draw conclusions. Advanced Analytics could include human-developed or machine-developed algorithms and encompasses, but is not limited to, the following: data mining, pattern and trend analysis, complex event processing, machine learning or deep learning, artificial intelligence, predictive analytics, big data analytics.

11. Does this Project, Program, or System use data to conduct electronic searches, queries, or analyses in an electronic database to discover or locate a predictive pattern or an anomaly indicative of terrorist or criminal activity on the part of any individual(s) (i.e., data mining)?[11] This does not include subject-based searches.	☒ No. ☐ Yes. If yes, please elaborate: *Click here to enter text.*
11(a) Is information used for research, statistical, or other similar purposes? If so, how will the information be de-identified, aggregated, or otherwise privacy-protected?	☒ No. ☐ Yes. If yes, please elaborate: *Click here to enter text.*

12. Does the planned effort include any interaction or intervention with human subjects[12] via pilot studies, exercises, focus groups, surveys, equipment or technology, observation of public behavior, review of data sets, etc. for <u>research purposes</u>	☒ No. ☐ Yes. If yes, please reach out to the DHS Compliance Assurance Program Office (CAPO) for <u>independent</u> review and approval of this effort.[13]

13. Does the Project, Program, or System provide role-based or additional privacy training for personnel who have access, <u>in addition</u> to annual privacy training required of all DHS personnel?	☒ No. ☐ Yes. If yes, please list: *Click here to enter text.*

[11] Is this a program involving pattern-based queries, searches, or other analyses of one or more electronic databases, where—

(A) a department or agency of the Federal Government, or a non-Federal entity acting on behalf of the Federal Government, is conducting the queries, searches, or other analyses to discover or locate a predictive pattern or anomaly indicative of terrorist or criminal activity on the part of any individual or individuals;

(B) the queries, searches, or other analyses are not subject-based and do not use personal identifiers of a specific individual, or inputs associated with a specific individual or group of individuals, to retrieve information from the database or databases; and

(C) the purpose of the queries, searches, or other analyses is not solely—

(i) the detection of fraud, waste, or abuse in a Government agency or program; or

(ii) the security of a Government computer system.

[12] Human subject means a living individual about whom an investigator conducting research: (1) obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens; or (2) obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens.

[13] For more information about CAPO and their points of contact, please see: https://www.dhs.gov/publication/capo or https://collaborate.st.dhs.gov/orgs/STCSSites/SitePages/Home.aspx?orgid=36. For more information about the protection of human subjects, please see DHS Directive 026-04: https://www.dhs.gov/sites/default/files/publications/mgmt/general-science-and-innovation/mgmt-dir_026-04-protection-of-human-subjects_revision-01.pdf.

14. Is there a FIPS 199 determination?[14]	☒ No. ☐ Yes. Please indicate the determinations for each of the following: Confidentiality: ☐ Low ☐ Moderate ☐ High ☐ Undefined Integrity: ☐ Low ☐ Moderate ☐ High ☐ Undefined Availability: ☐ Low ☐ Moderate ☐ High ☐ Undefined

PRIVACY THRESHOLD REVIEW

(TO BE COMPLETED BY COMPONENT PRIVACY OFFICE)

Component Privacy Office Reviewer:	(b) (6) (b) (7) (c)
PRIVCATS ID Number:	0016681
Date submitted to Component Privacy Office:	March 12, 2024
Concurrence from other Component Reviewers involved (if applicable):	N/A
Date submitted to DHS Privacy Office:	March 13, 2024

Component Privacy Office Recommendation:
Please include recommendation below, including what new privacy compliance documentation is needed, as well as any specific privacy risks/mitigations, as necessary.

(b) (5)

[14] 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. For more information, see https://www.nist.gov/itl/fips-general-information.

(b) (5)

(TO BE COMPLETED BY THE DHS PRIVACY OFFICE)

DHS Privacy Office Reviewer:	(b) (6)
DHS Privacy Office Approver (if applicable):	Click here to enter text.
PRIVCATS ID Number:	**0016681**
Date adjudicated by DHS Privacy Office:	March 13, 2024
PTA Expiration Date:	February 20, 2025

DESIGNATION

Privacy Sensitive System:	No
Category of System:	Pilot If "other" is selected, please describe: *Click here to enter text.*
Determination:	☒ Project, Program, System in compliance with full coverage. ☐ Project, Program, System in compliance with interim coverage. ☐ Project, Program, System in compliance until changes implemented. ☐ Project, Program, System not in compliance.
PIA:	Choose an item. *Click here to enter text.*
SORN:	Choose an item. *Click here to enter text.*

DHS Privacy Office Comments:
Please describe rationale for privacy compliance determination above, and any further action(s) that must be taken by Component.

(b) (5)

(b) (5)