epic.org

Electronic Privacy Information Center
1519 New Hampshire Avenue NW
Washington, DC 20036, USA

☎ +1 202 483 1140
📠 +1 202 483 1248
🐦 @EPICPrivacy
🌐 https://epic.org

COMMENTS OF THE ELECTRONIC PRIVACY INFORMATION CENTER

to the

Homeland Security Department

Agency Information Collection Activities: Public Perceptions of Emerging Technologies

86 Fed. Reg. 26228 / Docket No: DHS-2021-0015

July 12, 2021

The Electronic Privacy Information Center (EPIC) submits these comments in response to the Homeland Security Department's (DHS) proposed information collection, Public Perceptions of Emerging Technologies.[1] In this document DHS proposes a limited poll to collect information on Americans' perceptions of advanced surveillance technologies including facial recognition and AI powered systems.

EPIC urges DHS to go beyond a simple household poll and meaningfully account for the impacts of facial recognition and AI systems. EPIC urges DHS to 1) cease using facial recognition and AI-based technology in light of the serious threat these systems pose due to systemic problems with bias, accuracy, transparency and the disparate impacts they create or in the alternative to 2) perform rigorous impact assessments before implementing these technologies following the guidance outlined in part II.

EPIC is a public interest research center in Washington, D.C. EPIC was established in 1994 to focus public attention on emerging privacy and related human rights issues, and to protect

[1] 86 Fed. Reg. 26228.

privacy, the First Amendment, and constitutional values. EPIC has a particular interest in preserving the Privacy Act safeguards enacted by Congress.[2] EPIC also has an ongoing interest in DHS's use of machine learning and artificial intelligence on the agency's databases as well as DHS's use of biometrics, particularly the use of facial recognition as part of the Biometric Entry/Exit program.[3]

[2] *See, e.g.*, Comments of EPIC to the Department of Homeland Security, Correspondence Records Modified System of Records Notice, Docket No. DHS-2011-0094 (Dec. 23, 2011), http://epic.org/privacy/1974act/EPIC-SORN-Comments-FINAL.pdf; Comments of EPIC to the Department of Homeland Security, 001 National Infrastructure Coordinating Center Records System of Records Notice and Notice of Proposed Rulemaking, Docket Nos. DHS-2010-0086, DHS-2010-0085 (Dec. 15, 2010), http://epic.org/privacy/fusion/EPIC_re_DHS-2010-0086_0085.pdf; Comments of EPIC to the Department of Homeland Security, Terrorist Screening Database System of Records Notice and Notice of Proposed Rulemaking, Docket Nos. DHS-2016-0002, DHS-2016-0001 (Feb. 22, 2016), https://epic.org/apa/comments/EPIC-Comments-DHS-TSD-SORN-Exemptions-2016.pdf.

[3] *See e.g.*, Comments of EPIC to the Transportation Security Administration, Intent to Request Revision of Agency Information Collection Activity Under OMB Review: TSA PreCheck, Docket ID: TSA-2013-0001 (June 22, 2020), https://epic.org/apa/comments/EPIC-TSA-PreCheck-FRT-Comment-June2020.pdf; Comments of EPIC to the Department of Homeland Security, Agency Information Collection Activities: Biometric Identity, Docket No. 1651-0138 (Jul. 24, 2018), https://epic.org/apa/comments/EPIC-CBP-Vehicular-Biometric-Entry-Exit-Program.pdf; EPIC v. CBP (Biometric Entry/Exit Program), https://epic.org/foia/dhs/cbp/biometric-entry-exit/default.html (EPIC obtained a report which evaluated iris imaging and facial recognition scans for border control); EPIC Statement to U.S. House Committee on Homeland Security, "Border Security, Commerce and Travel: Commissioner McAleenan's Vision for the Future of CBP" (Apr. 24, 2018), https://epic.org/testimony/congress/EPIC-HHSC-CBP-Apr2018.pdf; Comments of EPIC to the Department of Homeland Security, Privacy Act of 1974: Implementation of Exemptions; Department of Homeland Security/U.S. Citizenship and Immigration Services—018 Immigration Biometric and Background Check (IBBC) System of Records, Docket Nos. DHS-2018-0002 and DHS-2018-0003 (Aug. 30, 2018) https://epic.org/apa/comments/EPIC-DHS-Immigration-Biometric-Database.pdf, Comments of EPIC to the Department of Homeland Security, Privacy Act of 1974: Implementation of Exemptions; Department of Homeland Security (DHS)/U.S. Customs and Border Protection (CBP)-024 CBP Intelligence Records System (CIRS) System of Records, Docket Nos. DHS-2017-0026 and 0027 (Oct. 23, 2017), https://epic.org/apa/comments/EPIC-CBP-Intelligence-Records-System-Comments.pdf, Comments of EPIC to the Department of Homeland Security, Privacy Act of 1974: Implementation of Exemptions; Department of Homeland Security/U.S. Immigration and Customs Enforcement-016 FALCON Search and Analysis System of Records, Docket Nos. DHS-2017-0001 and 2017-0002 (June 5, 2017), https://www2.epic.org/apa/comments/EPIC-DHS-FALCON-Database-Comments.pdf, EPIC v. CBP (Analytical Framework for Intelligence), https://www2.epic.org/foia/dhs/cbp/afi/ (EPIC obtained training materials showing that Palantir provides the infrastructure for the FALCON database).

I. DHS Programs Involving Artificial Intelligence and Facial Recognition Raise Issues of Privacy, Security, and Bias.

DHS already has several programs that use facial recognition and/or artificial intelligence despite the numerous privacy, security, and bias issues associated with the technologies. A few of those programs and some of their associated issues are described below.

a. Biometric Entry/Exit Program

Congress authorized DHS to develop a biometric entry and exit system in 2004, and then transferred this responsibility to the U.S. Customs and Border Protection (CBP) in 2013.[4] The use of facial recognition was not specified by Congress and Congress only authorized the biometric entry-exit program for non-U.S. citizens. CBP began testing biometric identity verification at select airports and land ports in 2016 and then operationalized it in 2017 through its Traveler Verification Service (TVS).[5] More than 23 million international travelers were involved in CBP's facial biometrics process in Fiscal Year 2020.[6] Currently, CBP has deployed facial comparison technology at 196 airports for identity verification at entry and 29 airports and 11 seaports for verification at exit.[7]

CBP's biometric entry and exit process through TVS uses facial recognition to verify travelers' identities when they enter or exit the U.S. by air, sea, or land.[8] Before travelers arrive,

[4] Biometric Air Exit Standard Operating Procedure, U.S. Customs and Border Protection (March 2019), https://epic.org/foia/dhs/cbp/biometric-entry-exit-alt-screening-procedures/Biometric-Air-Exit-SOP-Mar2019.pdf.
[5] Privacy Impact Assessment for Traveler Verification Service DHS/CBP/PIA-056 (November 14, 2018), https://www.dhs.gov/sites/default/files/publications/privacy-pia-cbp056-tvs-january2020_0.pdf.
[6] CBP Year in Review: Agency Adapts to Secure and Facilitate Essential Trade and Travel amid Pandemic (Feb. 4, 2021), https://www.cbp.gov/newsroom/national-media-release/cbp-year-review-agency-adapts-secure-and-facilitate-essential-trade.
[7] CBP Biometrics, https://biometrics.cbp.gov/.
[8] DHS/CBP/PIA-056 Traveler Verification Service (Nov. 14, 2018) (hereinafter Traveler Verification Service PIA), https://www.dhs.gov/sites/default/files/publications/privacy-pia-cbp056-tvs-january2020_0.pdf.

CBP creates localized photo “galleries” for TVS to use for facial recognition.[9] CBP builds these galleries based on Advance Passenger Information System (APIS) data from carriers who send biographic information about incoming travelers.[10] When CBP does not have advance passenger information, CBP builds photo galleries based on lists of frequent travelers at that entry point.[11] To populate the galleries, CBP pulls from the Automated Targeting System Unified Passenger Module system (ATS-UPAX) to obtain photos from U.S. passports and visas from the State Department, previous CBP inspections, and other DHS actions.[12] TVS uses the gallery photographs to generate biometric templates.[13] CBP does not allow travelers, including U.S. citizens to opt-out of having their photos used in the facial recognition galleries used by TVS.

CBP collects biometric information from passengers by taking real-time photos as they enter or exit the country.[14] CBP partners with private airlines, airport authorities, cruise lines, and other government agencies to take photos of travelers with cameras and transmit the photos to CBP’s TVS cloud.[15] Once TVS receives a photo, it compares the photo against the biometric template to produce a “match” or “no-match” result.[16]

[9] Traveler Verification Service PIA at 5.
[10] *Id*.
[11] *Id*.
[12] *Id* at 39.
[13] Traveler Verification Service PIA at 6.
[14] *Id* at 6.
[15] *Id*.
[16] DHS Data Privacy and Integrity Advisory Council, Report 2019-01, Privacy Recommendations in Connection with the Use of Facial Recognition Technology at 3 (Feb. 26, 2019) (hereinafter “DPIAC Facial Recognition Recommendations”), https://www.dhs.gov/sites/default/files/publications/Report%202019-01_Use%20of%20Facial%20Recognition%20Technology_02%2026%202019.pdf.

i. CBP does not perform necessary audits of private partners and subcontractors to ensure facial recognition images are secure and used only for legitimate purposes.

Despite the risks of using highly sensitive facial images for facial recognition, CBP has failed to ensure that its database of biometric information is secure.[17] The DHS has classified biometric information as sensitive personal identifying information, recognizing that lost or compromised biometric information can cause substantial harm, embarrassment, inconvenience, or unfairness to victims of a breach.[18] The Government Accountability Office (GAO) has also affirmed that the consequences of a breach of facial images may be significantly higher than for other information because faces are unique, permanent, and unchangeable.[19]

Despite the sensitivity of biometric data, in one breach CBP allowed approximately 184,000 facial recognition images to be exposed and 19 images to be posted to the dark web.[20] During a pilot of facial recognition technology in 2019, one of CBP's subcontractors downloaded biometric data and stored it on a separate network, which was subsequently hacked.[21] After investigating the breach, the DHS Office of Inspector General (OIG) found that CBP failed to implement all available IT security controls, including an acknowledged best practice. The OIG concluded that the CBP did not fulfill its IT security responsibilities.[22] CBP's failure to prevent a breach demonstrates that the agency is unable to safeguard sensitive biometric data.

[17] Joseph Cuffari, Review of CBP's Major Cybersecurity Incident During a 2019 Biometric Pilot, Dep't of Homeland Sec. Off. of Inspector Gen. (Sept. 21, 2020) (hereinafter OIG Review of 2019 Biometric Pilot), https://www.oig.dhs.gov/sites/default/files/assets/2020-09/OIG-20-71-Sep20.pdf.
[18] *Id* at 4.
[19] U.S. Gov't Accountability Off., GAO-20-522 Facial Recognition Technology: Privacy and Accuracy Issues Related to Commercial Uses (Jul. 2020), https://www.gao.gov/assets/gao-20-522.pdf.
[20] OIG Review of 2019 Biometric Pilot at 6.
[21] *Id*.
[22] *Id* at 12.

CBP also partners with commercial airlines and cruise lines for identity verification but has not audited most of them.[23] CBP's first audit of a commercial partner occurred in March 2020, three years after CBP began partnerships with private carriers.[24] As of May 2020, CBP had only audited one of its 20 airline partners and did not have a plan to audit others.[25] CBP also prohibits partners from retaining photos or using them for business purposes, but acknowledged that partners may sometimes capture photos, store them on their own IT infrastructure, and use them for business purposes separate from TVS.[26] The dearth of audits demonstrates CBP's lack of appreciation for the risks posed by facial recognition.

ii. CBP's facial recognition system may exhibit racial bias to produce less accurate results for people of non-white and non-male faces.

CBP's facial recognition programs risk misidentifying travelers in a biased manner because facial recognition algorithms are generally less accurate at identifying racial minorities.[27] In 2019, National Institute of Standards and Technology (NIST) published a report finding that the accuracy of facial recognitions can vary significantly depending on race, gender, and age.[28] With high quality images, false positives were highest among African and Asian people and lowest among Eastern European people.[29] Using lower-quality border crossing images, false negatives were higher for people born in Africa and the Caribbean, with a stronger effect for older individuals.[30] Given that CBP's facial recognition system focuses on identifying foreign nationals to return quick "match" or

[23] U.S. Gov't Accountability Off., GAO-20-568 Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues at 20 (September 2020) (hereinafter GAO Facial Recognition Report), https://www.gao.gov/products/GAO-20-568.
[24] *Id*.
[25] *Id*.
[26] *Id* at 49.
[27] Patrick Grother, Mei Ngan, and Kayee Hanaoka, NISTIR8280 Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects (Dec. 2019), https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf.
[28] *Id*.
[29] *Id*.
[30] *Id*.

"no match" results, CBP's system risks misidentifying people.[31] Misidentification at a border crossing can subject individuals to excessive surveillance and invasive searches, and may risk wrongful denial of entry to the US. Although CBP has touted its TVS algorithm to have a 97% accuracy rate, it may still misidentify a significant number of travelers due to the volume of travelers moving through U.S. airports, seaports, and land ports. In addition, CBP states that its priority mission is to safeguard America's borders, "thereby protecting the public from dangerous people" while facilitating legitimate trade and travel.[32] If CBP officials or partners assume that foreign visitors, immigrants, and refugees are dangerous, that may exacerbate and reinforce the effect of any algorithmic bias that leads to incorrect matches.

b. FALCON Search & Analysis System

The DHS/ICD-016 FALCON Search & Analysis System (FALCON-SA) enables U.S. Immigration and Customs Enforcement (ICE) personnel to "store, search, analyze, and visualize volumes of existing information in support of ICE's mission to enforce and investigate violations of U.S. criminal, civil, and administrative laws."[33] FALCON-SA was originally made available to ICE's Homeland Security Investigations (HSI) division agents, criminal research specialists, and analysts in HSI offices in the U.S. and abroad.[34] Access could also be granted to personnel "assigned to ICE HSI or to an ICE HSI-led task force," as well as "HSI contractors supporting a specific HSI

[31] *Id*.

[32] U.S. Customs and Border Prot., Biometric Air Exit: Standard Operating Procedure (Mar. 2019), https://epic.org/foia/dhs/cbp/biometric-entry-exit-alt-screening-procedures/BiometricAir-Exit-SOP-Mar2019.pdf.

[33] *Privacy Impact Assessment Update for the FALCON Search & Analysis System*, DHS 1, Oct. 11, 2016 (hereinafter "FALCON-SA PIA"), https://www.dhs.gov/sites/default/files/publications/privacy_pia_ice_falconsa_january2014.pdf.

[34] *Id.* at 7.

mission with a documented need to know."[35] Earlier this year, ICE said it would be updating the Privacy Impact Assessment (PIA) to limit the use of FALCON-SA to only personnel within HSI.[36]

FALCON-SA data consists of records routinely ingested from several DHS databases, as well as *ad hoc* ingestions that are manually entered by authorized FALCON-SA users.[37] Data is routinely ingested (usually at least once every 48 hours) from the following Privacy Act Systems of Records:[38]

- DHS/ICE-006 ICE Intelligence Records System (IIRS) SORN and DHS/ICE-009 External Investigations SORN (including law enforcement and intelligence reports, and reports of suspicious activity or threats);
- DHS/ICE-008 Search, Arrest, and Seizure Records SORN and DHS/CBP-013 Seized Assets and Case Tracking System (SEACATS) SORN (including ICE and Customs and Border Protection fines, penalties, and seized goods records);
- DHS/ICE/PIA-045 Investigative Case Management PIA and DHS/ICE-009 External Investigations SORN (including border screening records and law enforcement investigation data);
- DHS/ICE-011 Immigration and Enforcement Operational Records System (ENFORCE) SORN (including information about victims and witnesses of human trafficking and records documenting arrest, detention, and removal);
- FALCON Tipline (FALCON-TL) (including reports of suspicious or suspected illegal activity); and

[35] *Id.*
[36] Department of Homeland Security, *DHS/ICE/PIA-032 FALCON Search & Analysis System (FALCON-SA)* (Jan. 22, 2021), https://www.dhs.gov/publication/dhsicepia-032a-falcon-search-analysis-system-falcon.
[37] FALCON-SA PIA at 4.
[38] *Id.* at 32-34.

- FALCON Data Analysis & Research for Trade Transparency System (FALCONDARTTS) (including the Specially Designated Nationals List and other foreign asset information).

FALCON-SA data contains a variety of personally identifiable information (PII), including name and date of birth, citizenship and immigration data, border crossing data, criminal history and associates, contact information, family relationships, photographs and other media, and employment and education information.[39] FALCON-SA covers not only those who violated the law, but also suspected violators and individuals associated with an ICE investigation.[40] FALCON-SA thus includes information about witnesses, victims, sources, law enforcement personnel, and others.[41]

FALCON-SA provides users with a searchable index, from which they may analyze and share data.[42] Analytical results include maps, charts, and other visual outputs that "allow the user to identify links or connections that may have been previously unknown."[43] Users may save these analytical results in their private FALCON-SA workspace, create reports, share results with other users, or perform other tasks.[44] ICE personnel have expressed surprise about the power of FALCON-SA's analytical capabilities to link disparate data points.[45]

[39] *Notice of a new Privacy Act system of records*, 82 Fed. Reg. 20,905 (May 4, 2017) (hereinafter "FALCON-SA SORN").
[40] *See* FALCON-SA SORN at 20, 907.
[41] *Id.*
[42] FALCON-SA PIA at 3.
[43] *Id.*
[44] *Id.*
[45] *See* George Joseph, *Data Company Directly Powers Immigration Raids in Workplace*, WYNC (Jul. 16, 2019), https://www.wnyc.org/story/palantir-directly-powers-ice-workplace-raids-emails-show/ (quoting a former ICE HSI special agent stating, "It was just amazing how stuff would get linked by this phone number, by this address. And not only linking, but it would show you who or what is at the center of all that").

i. The FALCON database is used to identify and target vulnerable immigrant populations.

FALCON-SA has been used to support controversial ICE operations such as workplace raids.[46] In January of 2018, ICE agents used the FALCON mobile application to share information with the agency command center during an orchestrated workplace raid of almost 100 7-Eleven locations across 17 states.[47] In August of 2019, ICE used FALCON-TL software to log tips pertaining to alleged illegal hiring and employment activity at Koch Foods.[48] Those tips are routinely ingested into FALCON-SA and are then linked to other information in the database.[49] The resulting workplace raid led to the arrest of 680 migrant workers.[50]

ii. The FALCON database is prone to data errors, contains illegally obtained data, risks privacy through data breaches, and fails to meet basic standards of transparency.

ICE has asserted that FALCON-SA "assists the human evaluation and decision making process and helps reduce human error and analytic uncertainty by presenting information already available to the user in a common sense fashion."[51] Yet both routine and *ad hoc* data transfers create the potential for flawed data to amplify human error, not reduce it. Social media data may be uploaded to FALCON-SA either through routine ingests from ICE's Intelligence Records System

[46] *See* Douglas MacMillan & Elizabeth Dwoskin, *The war inside Palantir: Data-mining firm's ties to ICE under attack by employees*, WASH. POST (Aug. 8, 2019), https://www.washingtonpost.com/business/2019/08/22/war-inside-palantir-data-mining-firms-ties-ice-under-attack-by-employees/; *Failing to Do Right: The Urgent Need for Palantir to Respect Human* Rights, AMNESTY INT'L (Sept. 28, 2019), https://www.amnestyusa.org/wp-content/uploads/2020/09/Amnest-International-Palantir-Briefing-Report-092520_Final.pdf.

[47] Joseph, *supra* note 45; *see also* Patricia Mazzei, *Immigration Agents Target 7-Eleven Stores in Push to Punish Employers*, N.Y. TIMES (Jan. 10, 2018), https://www.nytimes.com/2018/01/10/us/7-eleven-raids-ice.html.

[48] *Breaking: Palantir's Technology Used in Mississippi Raids Where 680 Were Arrested*, MIJENTE (Oct. 4, 2019), https://mijente.net/2019/10/palantirpowersraids/.

[49] *Id.*

[50] MacMillan & Dwoskin, *supra* note 46.

[51] FALCON-SA PIA at 2.

(IIRS) or as *ad hoc* additions by ICE agents.[52] ICE personnel may then use FALCON-SA to conduct trend analysis or link and network analysis, both of which are intended to gain insight into a subject's current or future activities.[53] The accuracy of this data is not guaranteed, particularly as IIRS records may be pulled from commercial databases.[54] Data accuracy issues may also arise when gang database information is ingested as a result of ICE partnerships with local and state law enforcement agencies.[55] Gang databases are notoriously flawed, and individuals may appear in the datasets as a result of officer bias, ungrounded accusations, or plain error.[56]

Allowing flawed data to be uploaded to and analyzed by FALCON-SA leads to wrongful arrest and unjustified surveillance. FALCON-SA has been exempted from portions of the Privacy Act that could mitigate negative effects. For example, the FALCON-SA database is exempted from subsection (e)(1) of the Privacy Act that requires information to be relevant and necessary to be placed in a database.[57] Given the vast amount of data contained in FALCON-SA, it is highly likely that inaccurate, and thus irrelevant, data is used for analytical processes.

Documents disclosed by ICE suggest some sources of FALCON-SA data may have been obtained illegally. The FALCON-SA database at one time had access to data derived from Black Asphalt, a private intelligence database used by police to share reports, including PII, about

[52] *See* Faiza Patel et al., *Social Media Monitoring: How the Department of Homeland Security Uses Digital Data in the Name of National Security*, BRENNAN CTR. FOR JUST. 28 (Mar. 11, 2020), https://www.brennancenter.org/sites/default/files/2020-03/SocialMediaMonitoring.pdf (explaining how routine ingests from the IIRS include social media data collected by mobile device extraction tools).

[53] *See id.*; *Social Network Analysis Advanced Reference Guide*, ICE 1, https://epic.org/foia/dhs/ice/palantir-databases/FALCON-Social-Network-Analysis-Reference-Guide.pdf.

[54] Patel, *supra* at note 52 at 6.

[55] Will Lennon, *The Major Blindspot Undermining Sanctuary Cities and Helping ICE*, THE APPEAL (Nov. 19, 2020), https://theappeal.org/politicalreport/data-pipelines-sanctuary-cities-ice/.

[56] *See id.* (discussing the potential for an individual to be added to a gang database when officers fake evidence or a "reliable informant" makes an accusation of gang affiliation).

[57] *Final Rule*, 84 Fed. Reg. 45,641 (Aug. 30, 2019) (hereinafter "FALCON-SA Final Rule").

American motorists.[58] The program fueled aggressive civil asset forfeiture operations by several police departments.[59] At least two states have prohibited law enforcement use of Black Asphalt based on their assessment that it may violate laws governing privacy and civil liberties.[60] FALCON-SA also had or maintains access to data from Cellebrite, a mobile device data extraction tool used by many law enforcement agencies, including CBP.[61] Following the Supreme Court's decision in *Riley v. California*, EPIC and other privacy advocates have challenged the use of Cellebrite to extract information from mobile devices as unconstitutional searches.[62]

FALCON-SA data retention policies—providing a retention period of up to 20 years—allow data ingested from Black Asphalt and Cellebrite to remain a factor in analytical processes despite questions of relevance and legality.[63] If they are not deleted by the user, analytical results and search queries that incorporate this data may be retained for up to 30 years.[64] ICE should not allow information that was collected in violation of the law to then serve as the basis for its analytical processes.

ICE has allowed a broad array of outside actors access to FALCON-SA and has failed to adequately address the potential for harmful data breaches. The initial user base for FALCON-SA

[58] *ICE TECH Modernization Program Operational Requirements Document*, DHS Office of the Chief Information Officer 12, Feb. 28, 2014; Spencer Woodman, *Palantir Provides the Engine for Donald Trump's Deportation Machine*, THE INTERCEPT (Mar. 2, 2017), https://theintercept.com/2017/03/02/palantir-provides-the-engine-for-donald-trumps-deportation-machine/; Michael Sallah et al., *Stop and seize*, Wash. Post (Sept. 6, 2014), https://www.washingtonpost.com/sf/investigative/2014/09/06/stop-and-seize/?utm_term=.dbc2ecc080e1.
[59] Sallah et al., *supra* note 58.
[60] *Id.*
[61] Woodman, *supra* note 58.
[62] *Riley v California*, 135 S.Ct. 2473 (2014) (holding that warrantless search of a cellphone is unconstitutional); *see* Letter from Elec. Priv. Info. Ctr. to Hon. Rand Paul & Hon. Gary Peters, U.S. Senate Comm. on Homeland Sec. (July 10, 2018), https://epic.org/testimony/congress/EPIC-HHSC-WarentlessSearchesBorder-July2018.pdf ("CBP and ICE are searching electronic devices without even reasonable suspicion despite the U.S. Supreme Court having recognized a Constitutionally significant privacy interest in mobile devices").
[63] FALCON-SA PIA at 22.
[64] *Id.*

was excessively broad, allowing access not only to HSI agents, but also personnel from other agencies who are affiliated with an HSI-led task force and HSI contractors.[65] Allowing so many users access to such a sensitive and extensive database creates the potential for security breaches and misuse. These dangers have already come into reality with respect to other technologies utilized by DHS. In 2019, the DHS OIG issued a report detailing a major privacy breach resulting from a cyber-attack on a subcontractor.[66] In that instance, sensitive biometric data was accessed by a malicious actor despite the subcontractor's commitment to protect PII from identity theft or misuse.[67] FALCON-SA presents similar concerns as the database contains a multitude of PII.

c. Analytic Framework for Intelligence

CBP uses several systems to automate the process of reviewing data in an attempt to identify relationships between individuals. CBP offices use the resulting intelligence products for law enforcement purposes, including investigations and prosecution. CBP's Office of Intelligence (OI) created the CBP Intelligence Records System (CIRS) to contain the information that is incorporated into intelligence products: this information includes OI's raw intelligence, public source information, and information initially collected by other CBP offices. CBP uses the Analytic Framework of Intelligence (AFI) and the Intelligence Reporting System (IRS) to analyze information and develop intelligence products. These products are then disseminated to CBP executive management, CBP units, other government agencies, and the "Intelligence Community."

Little information is available about AFI and even less is available about IRS, but the fact that both AFI and IRS are used in intelligence efforts raises concern over their deployment of artificial intelligence and algorithm-based analytical tools. Given the lack of transparency around

[65] FALCON-SA PIA at 7.

[66] DHS Office of the Inspector General, *Review of CBP's Major Cybersecurity Incident during a 2019 Biometric Pilot* at 5 (Sept. 21, 2020), https://www.oig.dhs.gov/sites/default/files/assets/2020-09/OIG-20-71-Sep20.pdf.

[67] *Id.*

how these systems use algorithms and AI, whether they are accurate, whether they protect privacy rights, and whether they have discriminatory impacts, CBP should stop using them.

i. AFI fails to meet baseline expectations of transparency, accuracy, and lacks privacy protections.

Very little is known about how AFI operates. EPIC litigated FOIA requests against CBP, and the subsequently released documents revealed Palantir's deep ties to AFI.[68] CBP stated in one document that both AFI and Palantir are authorized to store or process sensitive but unclassified data and information and that both AFI and Palantir data are accessible to AFI users.[69] Training documents obtained by EPIC list entire training sessions just called "Palantir."[70] In documents from CBP Office of Intelligence and Investigative Liaison asking for AFI Feedback, questions refer to "the AFI system and the Palantir Tool."[71] It's clear that Palantir,[72] an American technology company providing data-mining software, is a significant player in how AFI works. The software company is known for using artificial intelligence and machine learning algorithms analyze disparate sources of data, and its software is what likely drive AFI's search engine and analytic linking capabilities. The workings of Palantir's software remain a black box despite their application on vast amounts of data connected to AFI.

AFI is connected to numerous data sources and consequently has access to a broad amount of sensitive information. For example, AFI has access to data from CBP TECS, Border Crossing

[68] EPIC v. CBP (Analytical Framework for Intelligence, https://epic.org/foia/dhs/cbp/afi/.

[69] U.S. CUSTOMS AND BORDER PROTECTION, ANALYTICAL FRAMEWORK FOR INTELLIGENCE OPERATIONAL STATUS & SECURITY (obtained Apr. 8, 2014), https://epic.org/foia/dhs/cbp/afi/14-04-08-CBP-FOIA-20150205-Production-p4.pdf (obtained through EPIC's FOIA litigation against CBP in EPIC v. CBP, No. 14-1217 (D.D.C. 2016)).

[70] *See, e.g.*, U.S. CUSTOMS AND BORDER PROTECTION, ANALYTICAL FRAMEWORK FOR INTELLIGENCE RRB TRAINING (obtained Apr. 8, 2014), https://epic.org/foia/dhs/cbp/afi/14-04-08-CBP-FOIA-20150205-Production-p4.pdf (obtained through EPIC's FOIA litigation against CBP in EPIC v. CBP, No. 14-1217 (D.D.C. 2016)).

[71] *Id.*

[72] Palantir, https://www.palantir.com/.

Information (BCI), and the Automated Targeting System (ATS).[73] ATS itself allows users to search data across many different databases—for example, the FBI Terrorist Screening Database—and uses that data to create risk-based assessments of travelers, cargo, and conveyances.[74] Additionally, AFI has access to ICE intelligence products and legacy data from the National Security Entry Exit Registration System (NSEERS),[75] a controversial system from the Bush Administration requiring visa-holders from a list of countries, predominantly Muslim, to register with the federal government.[76] Furthermore, AFI now acts as a portal to Law Enforcement Information Sharing Services (LEISS) data sources, which give approved users access to state and local criminal repositories.[77]

The type of information accessible through these data systems includes biographical information, personal associations, travel itineraries, immigration records, home and work addresses, and physical traits like fingerprints, scars, or tattoos.[78] AFI also permits certain users to upload information from other sources, such as the Internet (including social media) or traditional news media.[79]

AFI uses the information from the various data sources to draw connections and make decisions that precipitate surveillance that have real-world consequences. Yet there is no transparency on how the underlying AI and algorithms work and consequently there is no public

[73] DHS/CBP/PIA-010(a) - Analytical Framework for Intelligence (AFI) at 3.
[74] *See* DHS/CBP/PIA-006(e) - Automated Targeting System at 1 (Jan. 17, 2017) (Add. updated May 5, 2021).
[75] *Id.* at 10.
[76] *See, e.g.*, Spencer Woodman, *Palantir Provides the Engine for Trump's Deportation Machine*, THE INTERCEPT, (Mar. 2, 2017), https://theintercept.com/2017/03/02/palantir-provides-the-engine-for-donald-trumps-deportation-machine/.
[77] DHS/CBP/PIA-010(a) - Analytical Framework for Intelligence (AFI) at 11. *See also* DHS/ICE/PIA-051 - Law Enforcement Information Sharing Service (LEIS Service) (June 28, 2019).
[78] Spencer Woodman, *Documents suggest Palantir could help power Trump's 'extreme vetting' of immigrants*, THE VERGE (Dec. 21, 2016), https://www.theverge.com/2016/12/21/14012534/palantir-peter-thiel-trump-immigrant-extreme-vetting.
[79] DHS/CBP/PIA-010(a) - Analytical Framework for Intelligence (AFI) at 4.

understanding of or accountability for how CBP makes decisions in targeting certain individuals or locations.

The lack of transparency and use of black box algorithms is exacerbated by the fact that AFI has been exempted from many of the protections of the Privacy Act. For example, AFI is exempted from the Privacy Act requirement to maintain records with accuracy, relevance, timeliness, and completeness as is "reasonably necessary" to assure fairness to the individual subject under §552a(e)(5).[80] The stakes of inaccurate information are high when such information is being used to investigate and prosecute civilians. With approval, certain non-CBP employees can gain access to AFI,[81]including the Immigration and Customs Enforcement's primary deportation arm, the office of Enforcement and Removal Operations.[82] DHS' own PIA acknowledges the risk of AFI containing erroneous information about individuals.[83]

The privacy risks created by AFI have increased since the system moved to an open-source format that now replicates and stores multiple copies of the source data. This highly distributed file system increases the risk of unauthorized access to multiple copies of AFI and source data sets.[84] DHS' own PIA cautions that there is a risk of unauthorized access.[85] Furthermore, even once employees are given access, AFI does not track users by Component Office or mission: there are thus few safeguards in place to ensure that a user accesses information only for narrow job purposes.[86] There is functionally no way to ensure that AFI is not abused for illegitimate purposes.

[80] AFI Privacy Act Exemption Final Rule at 47768; 5 U.S.C. § 552a(e)(5).
[81] DHS/CBP/PIA-010(a) - Analytical Framework for Intelligence (AFI) at 10.
[82] *See* Woodman, *Documents suggest Palantir could help power Trump's 'extreme vetting' of immigrants*, THE VERGE, *supra* note 115.
[83] DHS/CBP/PIA-010(a) - Analytical Framework for Intelligence (AFI) at 17.
[84] *Id.*
[85] *Id.* at 15.
[86] *Id.*

ii. AFI creates disparate impacts on marginalized communities.

The premise of AFI is that, based on collected data like associations, travel itineraries, proximity to certain locations, it can assess and predict which individuals warrant law enforcement attention. But the secrecy around how these assessments and predictions are created means there is no transparency on whether discriminatory factors like race, nationality, and religion, or proxies for these categories are being used.[87] This is especially concerning because AFI contains travel and immigration data, since Muslim travelers are consistently profiled and subject to additional scrutiny.[88] If AFI uses existing data on investigated individuals to make predictions on new associated individuals who should also be subject to scrutiny, it's entirely possible that Muslim travelers who have been investigated on the basis of their religion comprise existing data used to identify new suspicious travelers, most likely other Muslim travelers. The result is a cycle of discrimination that reinforces itself, where black box algorithms lend such discriminatory practices an air of legitimacy. It's also worth noting that many seemingly neutral factors, like home address or travel patterns, often act as proxies for race, ethnicity, or religion.

If Muslim and Latinx travelers are disproportionately represented in AFI data and disproportionately identified as potential risks, they will subsequently be disproportionately subject to law enforcement investigations and prosecution. Moreover, the above-mentioned concerns of accuracy, privacy, and misuse of data would all disproportionately apply to these communities.

[87] Faiza Patel, Rachel Levinson-Waldman, and Raya Koreh, *Social Media Monitoring*, Brennan Center for Justice (May 22, 2019) (updated Mar. 11, 2020), https://www.brennancenter.org/our-work/research-reports/social-media-monitoring#footnote149_6du5hbc.

[88] *See, e.g.*, Michael T. Luongo, *Traveling While Muslim Complicates Air Travel,* N.Y. TIMES (Nov. 7, 2016), https://www.nytimes.com/2016/11/08/business/traveling-while-muslim-complicates-air-travel.html.

d. Intelligence Reporting System

Very little is publicly available about IRS,[89] CBP's other intelligence information system cited alongside AFI in its CIRS SORN. It seems that IRS is used similarly to AFI in analyzing and producing intelligence information, as well as developing finished intelligence products.[90] It also seems that IRS may operate as a data source for AFI—IRS data is available through AFI, including CBP Field Information Reports, Homeland Security Intelligence Reports, and ICE NameTrace data.[91]

Although a Privacy Impact Assessment has been written for AFI, none seem to exist for IRS, though a 2017 SORN for CIRS claimed it was forthcoming.[92] Given the lack of information about what data is potentially collected, indexed, or analyzed in IRS, it is unclear but likely that the Privacy Act exemptions that apply to CIRS also apply to certain information in IRS.[93]

Without further information on IRS, it is impossible for the public to understand what information of theirs is being collected, stored, and processed. A PIA is necessary for transparency and accountability. Along with the type of information within IRS, the public also needs the information to understand what kind of algorithmic decision-making IRS deploys so as to identify accuracy and privacy issues. It's also unclear if IRS, like AFI, is a result of a CBP partnership with Palantir. If so, the same issues that arose with AFI regarding transparency and access also apply to IRS. As long as IRS is kept a black box operating system, there is also a chance that it is producing disparate impacts with no repercussions.

[89] 2017 CIRS SORN at 44199.

[90] *Id.*

[91] DHS/CBP/PIA-010(a) - Analytical Framework for Intelligence (AFI) at 8-9.

[92] 2017 CIRS SORN at 44199.

[93] *See Id. See also* CIRS Privacy Act Exemption Proposal.

II. DHS should perform meaningful impact assessments before developing and deploying any facial recognition system or other system using artificial intelligence technology.

Where FR/AI technologies are used, DHS should require meaningful impact assessments. As it stands, DHS's use of AI and facial recognition technologies is largely obscured from public view. Even when DHS does publish a PIA for a system or program, it often includes very little concrete information about how the AI works, whether DHS is tracking inaccuracies, discrimination, and privacy harms, and whether DHS has measures in place to correct those harms. If DHS continues using these technologies, it should require more comprehensive and rigorous impact assessments.

a. Principles underlying a meaningful algorithmic impact assessment.

For any impact assessment to be a useful endeavor it must be sufficiently detailed to identify real problems, preemptive so that an impact assessment can shape a proposed project, and ongoing so that DHS can understand how existing systems are performing. An impact assessment that cannot alter the substance of a proposed system is simply a box checking exercise. DHS should insert an impact assessment into their system development process so that the results of the assessment can be incorporated into design decisions. Ultimately, the results of an impact assessment should be weighted heavily when DHS decides whether to continue with an AI or FR system.

DHS should also stand up an independent office or work with an outside agency like NIST to perform impact assessments. If DHS stands up a new office or new task force within the Privacy Office, the agency must provide the necessary resources and mandate full engagement by agency components. Impact assessors should be looped in early on a project and given full access rights to test AI systems and gather data.

DHS should look to the Universal Guidelines for Artificial Intelligence (UGAI) as the theoretical framework underpinning its algorithmic impact assessments.[94] The UGAI set out 12 metrics that an AI system must meet: transparency; human determination; identification of AI system use; fairness; assessment and accountability; accuracy, reliability, and validity; data quality; public safety; cybersecurity; prohibition on secret profiling; prohibition on unitary scoring; and system termination when human control is no longer possible. Implementing these metrics requires a detailed assessment that considers both the technical performance of the system and its impact on individuals. DHS should also consider the GDPR's algorithmic impact assessment scheme,[95] Canada's impact assessment model,[96] and the proposed Algorithmic Accountability Act when developing an AI/FR assessment process.[97]

b. DHS's impact assessments must consider factors including bias, security, accuracy, accountability, transparency, data quality, and disparate impacts.

DHS's algorithmic impact assessments should include sections which detail each of the factors outlined below explaining the relevant technical details, data, and broader social concerns that underpin these factors. To make impact assessments meaningful, DHS must apply the factors, not simply provide boilerplate statements identifying concerns. DHS should also routinely update impact assessments as systems are implemented or expanded.

Systemic bias and accuracy should be primary concerns for DHS. Any DHS system should undergo rigorous testing to identify biases built into software. NIST's recent tests of facial recognition systems are a good model for DHS to follow in implementing bias testing.[98] Bias testing

[94] The Public Voice, *Universal Guidelines for Artificial Intelligence* (Oct. 23, 2018) [hereinafter *UGAI*], https://thepublicvoice.org/ai-universal-guidelines/.
[95] Article 35, General Data Protection Regulation, https://gdpr-info.eu/art-35-gdpr/.
[96] Algorithmic Impact Assessment Tool, Canada.ca https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html.
[97] H.R. 2231, 116th Cong. (2019 https://www.congress.gov/bill/116th-congress/house-bill/2231/text.
[98] Patrick Grother, Mei Ngan, and Kayee Hanaoka, NISTIR8280 Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects (Dec. 2019), https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf.

should be robust to cover a variety of possibly impacted groups including race/ethnicity, gender, age, income, citizenship status, religious affiliation, education or other factors. Because AI systems can be used on so many different types of data, DHS should choose to customize the factors tested based on the data used by an AI program and should justify that choice. Algorithmic assessors should also test systems for overall accuracy and set minimum required performance thresholds for AI systems. Accuracy standards should be context specific and set before accuracy testing is performed. These standards should be the highest where systems have the most consequential outcomes: law enforcement and immigration.

Security is another key factor that should be explicitly addressed in an algorithmic impact assessment. Security assessments should include both "hard" system security features like anti-hacking measures and "soft" features like data access permissions. The use of AI and facial recognition creates serious security concerns for those who are being surveilled or whose data is being collected, retained, and analyzed. DHS must ensure that its data use practices are secure enough to prevent harmful disclosures or opt not to collect and analyze sensitive data. Security includes protection from unauthorized users, but also from overly broad uses: implementing narrow use policies and giving any employees limited access for specific tasks will guard against dangerous uses of the data. Any security assessment should look closely at access rights and pay particular attention to government contractors or other third-party users. A security assessment should also consider the entity that designed an out-of-the-box system, asking whether that entity will continue to have access to data collected once DHS starts using it. The security section of an algorithmic impact assessment should identify potential shortfalls and detail required steps to prevent data breaches or other vulnerabilities. Updates to impact assessments should confirm DHS implemented the required steps to mitigate risks and look for new security vulnerabilities.

DHS should also mandate data quality testing to ensure that AI systems are built on data that is accurate and unbiased, and continue to evaluate data quality through the lifetime of a system. This analysis should identify the data sources that will be used in an AI system and detail any accuracy or bias concerns. A meaningful impact assessment will consider whether DHS should use a proposed data source, find an alternative, or cancel a project if there are no sufficiently reliable data sources available. Updates to an impact assessment should look at any new data sources used by a system or changes in existing data collection procedures.

Any impact assessment that DHS conducts on its AI and facial recognition tools must include an evaluation of disparate impacts. That evaluation should seek input from marginalized communities, provide explanation of systems in place to measure disparate impact, and contain a thorough description of mitigation and preventative efforts. To maintain meaningful safeguards against disparate impact, DHS must establish rigorous and objective processes for evaluating and terminating discriminatory technology. This requires detailed data on the populations effected by DHS's system to draw out deviations from expected performance metrics. An impact assessment should provide an explanation of how DHS defines the population impacted by its tool, a prediction for how a fair system would perform when used on that population, and an analysis of how DHS's system actually performs. DHS cannot simply identify risks, but must provide real metrics on a system's performance, identify the causes of disparate impacts, and require steps to resolve the disparate impact or dissolve the system.

An algorithmic impact assessment should enforce accountable systems that require human review of AI determinations and provide a mechanism for correcting errors. The impact assessment should consider what points in the system require human oversight and pay close attention to areas where information produced by AI is transmitted to other DHS systems or outside agencies. The assessment should also consider how easy it is to correct wrong information and whether there are greater barriers to resolving errors for some affected populations.

Finally, DHS should be radically transparent with its impact assessments. These assessments should be published without redactions and DHS should choose not to exempt AI systems from Privacy Act protections. DHS's liberal use of exemptions from the Privacy Act of 1974 has given substantial cover to its systems of records. Namely, DHS regularly exempts its systems from disclosure, access, and accuracy standards in the name of protecting law enforcement investigations from being compromised and of acknowledging the purported realities of factual ambiguities in such investigations. However, access and accuracy are even more important precisely in the law enforcement context, where the stakes of black box secrecy and inaccurate information can do irreparable harm to individuals and communities who are deprived of accountability tools. DHS itself has a strong incentive to be transparent in this area as inaccurate information produces inaccurate systems.

Conclusion

a. DHS should stop using facial recognition and AI-based surveillance technologies.

DHS should cease its use of AI-based surveillance and intelligence technologies in programs and databases like CIRS/AFI, FALCON-SA, and Biometric Entry/Exit. EPIC and 70 other organizations, including the American Civil Liberties Union and the Electronic Frontier Foundation, recently wrote a letter requesting that DHS stop its use of Clearview AI's facial recognition services, citing the privacy violations of its data scraping practices from social media and the secrecy that makes its accuracy unclear.[99] These same concerns—of intrusive social media surveillance and black box algorithms—apply to the AI and facial recognition technology implemented in CIRS/AFI,

[99] Coalition Letter to DHS Secretary Alexander Mayorkas re: Clearview AI (Apr. 19, 2021), https://justfutureslaw.org/wp-content/uploads/2021/04/Clearview-AI-sign-on-letter.pdf.

FALCON-SA, and Biometric Entry/Exit. Immediate suspension of these programs is also crucial to protecting the rights of people of color, who are likely disproportionately represented in intelligence report databases and disproportionately misidentified by biometric technology.

DHS should further end its relationship with and delete any data obtained from the controversial surveillance and analytics company Palantir. The risk that a private company like Palantir may have access to sensitive immigration information through AFI, and that CBP may have access to the troves of information that Palantir collects through its other work poses a serious and imminent threat to privacy and civil liberties.

b. DHS must perform rigorous and meaningful impact assessments before implementing any AI or facial recognition system.

EPIC urges DHS to develop and implement a rigorous algorithmic impact assessment process that consults with a variety of stakeholders and provides real insight to govern the development or maintenance of AI systems. Such a process can help protect the public from the most serious harms of automated decision making and black box data analysis. An impact assessment process can also help DHS develop more reliable systems and catch errors faster. Thank you for your time and consideration of EPIC's comments. Please contact Jeramie Scott at jscott@epic.org with any questions.

Respectfully Submitted,

Jeramie Scott
Jeramie Scott
EPIC Senior Counsel

Jake Wiener
Jake Wiener
EPIC Law Fellow

Bridget Amoko
Bridget Amoko
EPIC Clerk

Soojin Jeong
Soojin Jeong
EPIC Clerk

Nicole Mo
Nicole Mo
EPIC Clerk