March 8, 2018

Ranking Member Thompson and members of the House Committee on Homeland Security
United States House of Representatives

Re: DHS Visa Lifecycle Vetting Initiative

Ranking Member Thompson, Rep. Jackson Lee, Rep. Langevin, Rep. Richmond, Rep. Keating, Rep. Payne, Rep. Vela, Rep. Coleman, Rep. Rice, Rep. Correa, Rep. Demings, and Rep. Barragan,

On behalf of the Center for Democracy & Technology (CDT), I am writing to urge you to stop the Department of Homeland Security (DHS) from using unreliable and discriminatory social media monitoring techniques to deny immigration benefits and generate deportation leads. DHS is currently seeking a vendor to engage in collection and machine-learning analysis of social media posts to automatically flag 10,000 people per year—inside the country and abroad—for deportation investigations and denial of immigration benefits (such as visas).[1] This program, formerly known as the "Extreme Vetting Initiative," is a serious threat to civil rights, privacy, and free speech. I urge you to do everything in your power to stop this ill-considered effort.

In July 2017, DHS held an industry day to pursue a contract to automate the current manual process of screening applicants for entry to the United States and immigration benefits as well as the process of creating deportation leads.[2] In February 2018, DHS announced that it is moving forward with this contract following a closed procurement process.[3] The contractor would be responsible for "exploit[ing]" information from sources such as social media websites, blogs, conferences, and academic websites.[4] The contractor would be required to use automated processes to generate 10,000 investigative leads annually, to evaluate applicants' "probability of becoming a positively contributing member of society as well as their ability to contribute to the national interests," and to "assess whether an applicant intends to commit criminal or terrorist acts after entering the United States."[5] At the industry day, DHS expressed its interest in acquiring "machine learning capabilities" to do

---

[1] *See* Immigration & Customs Enforcement Homeland Security Investigations (hereinafter "ICE-HSI"), "ICE-HSI Data Analysis Service: Solicitation Number HSCEMD-17-R-0010," FedBizOpps.Gov, June 12, 2017. *See also* Brennan Center for Justice, ICE Extreme Vetting Initiative: A Resource Page, https://www.brennancenter.org/analysis/ice-extreme-vetting-initiative-resource-page.

[2] *See* ICE-HSI, "Extreme Vetting Initiative: STATEMENT OF OBJECTIVES (SOO)," June 12, 2017, available at FedBizOpps.Gov (hereinafter "Statement of Objectives").

[3] ICE-HSI, updated Feb. 20, 2018, https://www.fbo.gov/?s=opportunity&mode=form&tab=core&id=bd5d7ab15098274fe01b8e1daf37f0bb&_cview=1 (All future procurement actions related to this acquisition may be followed at the DHS Procurement system, Acquisition Planning Forecast System (APFS).); APFS forecast no. F2018040916, available at http://apfs.dhs.gov/.

[4] ICE-HSI, "Background," June 12, 2017, available at FedBizOpps.Gov.

[5] *Id.*

this vetting.[6] This initiative is now known as "Visa Lifecycle Vetting."[7] DHS's proposal is technologically infeasible, discriminatory by design, and likely to chill free speech.

DHS's proposal to automatically vet immigration applicants has no basis in law or in technological capabilities. The outcomes DHS intends to predict—whether a person will positively contribute to society or to the national interest—have no objective or legal definitions that can be used to train a reliable algorithm.[8] As CDT has explained in a recent white paper, automated tools for analyzing the text of social media posts cannot reliably interpret the meaning of a post or the speaker's intent.[9] In November 2017, 54 of the nation's leading computer science experts sent a letter to DHS stating that "no computational methods can provide reliable or objective assessments of the traits that ICE seeks to measure."[10] Any automated tools developed for making these determinations would be inherently inaccurate and biased.[11]

The use of automated tools to predict amorphous criteria would disproportionately harm minority groups. Machine learning models tend to reflect and even amplify any biases in the data used to train them. When these models are used for social media analysis, they often produce higher error rates when analyzing the speech of minority group members. For example, research shows that many popular social media analysis tools cannot accurately process tweets using language common among African-American Twitter users.[12] The tools in the study misidentified these tweets as not being in English at all, with one tool classifying the tweets as being in Danish with 99.9% accuracy.[13] These biases are exacerbated by the fact that most available tools for performing this kind of analysis only work on English-language text.[14]

DHS's proposal to exploit social media would violate companies' policies and users' privacy. As described, the Visa Lifecycle Vetting program relies on DHS's or its contractor's ability to access and analyze large amounts of

---

[6] ICE-HSI, "Q&A from Vendor Emails," July 27, 2017, available at FedBizOpps.Gov.

[7] *See* APFS forecast no. F2018040916, available at http://apfs.dhs.gov/.

[8] These phrases only appear in President Trump's original travel ban executive order, which has since been rescinded. See Exec. Order 13,769 (Jan. 27, 2017) at sec. 4.

[9] Natasha Duarte, Emma Llanso & Anna Loup, Center for Democracy & Technology, Mixed Messages? The Limits of Automated Social Media Content Analysis (2017), https://cdt.org/insight/mixed-messages-the-limits-of-automated-social-media-content-analysis/.

[10] Letter from 54 Technology Experts to Hon. Elaine Duke, Sec'y of Homeland Security, Nov. 16, 2017, available at https://www.brennancenter.org/sites/default/files/Technology%20Experts%20Letter%20to%20DHS%20Opposing%20the%20Extreme%20Vetting%20Initiative%20-%2011.15.17.pdf.

[11] For more details about why this program would be technically infeasible and discriminatory, *see* Natasha Duarte, *Automated "Extreme Vetting" Won't Work and Will Be Discriminatory*, Center for Democracy & Technology (Nov. 16, 2017), https://cdt.org/blog/automated-extreme-vetting-wont-work-and-will-be-discriminatory/.

[12] Su Lin Blodgett & Brendan O'Connor, *Racial Disparity in Natural Language Processing: A Case Study of Social Media African-American English* at 1-2, Proceedings of the Fairness, Accountability, and Transparency in Machine Learning Conference (2017), https://arxiv.org/pdf/1707.00061.pdf.

[13] *Id.*

[14] *See* Duarte et al., *supra* note 9, at 14–15.

1401 K Street NW, Suite 200 Washington, DC 20005

data from social media platforms. Many platforms have prohibited this type of access to their users' information even if it appears on a "public" profile, either through technical means or through explicit policies.[15] DHS has indicated its intent to use technical "workaround[s]" to evade these limitations. Since the program is limited to "publicly available" information, the need for workarounds indicates DHS's intent to collect and analyze social media information in bulk, rather than limiting its analyses to the individual targets of investigations. This type of surveillance would inevitably sweep in individuals who are not the targets of investigations, including US citizens and lawful permanent residents.

This initiative would chill free speech and association.The knowledge that the government will be regularly scouring online statements to make admission or deportation determinations will unquestionably pressure both visa applicants and recipients – and the people with whom they communicate – to censor themselves online. Visa applicants and holders are likely to self-censor or delete accounts with consequences for personal, business and travel-related activity. These risks are particularly acute in light of existing initiatives to ask travelers to identify all of their social media handles in order to obtain permission to travel to the United States.[16] Moreover, because of the continuous nature of this proposed social media monitoring, the resulting chill to free expression and association would be ongoing and not limited to the time frame during which people are applying for immigration benefits.

In November, CDT joined more than 100 civil society organizations and technical experts to urge the Department of Homeland Security to stop this program.[17] We ask you to use all of the tools at your disposal to stop this discriminatory, dangerous, and ineffective program from becoming a reality.

Sincerely,

Natasha Duarte, Policy Analyst
Center for Democracy & Technology
1401 K Street NW Suite 200
Washington, DC 20005

---

[15] *See, e.g.*, Lilly Hay Newman, *Facebook's Big "First Step" to Crack Down on Surveillance*, Wired (March 17, 2017), https://www.wired.com/2017/03/facebooks-big-first-step-crack-surveillance/; Chris Moody, *Developer Policies to Protect People's Voices on Twitter*, Twitter Developer Blog (Nov. 22, 2016), https://blog.twitter.com/developer/en_us/topics/community/2016/developer-policies-to-protect-peoples-voices-on-twitter.html.

[16] *See, e.g.*, Center for Democracy & Technology, *DHS Proposal to Ask for Social Media Identifiers a Threat to Privacy & Free Speech* (Aug. 22, 2016), https://cdt.org/press/dhs-proposal-to-ask-for-social-media-identifiers-a-threat-to-privacy-free-speech/.

[17] Letter from 56 organizations to the Department of Homeland Security, Nov. 16, 2017, https://www.brennancenter.org/sites/default/files/Coalition%20Letter%20to%20DHS%20Opposing%20the%20Extreme%20Vetting%20Initiative%20-%2011.15.17.pdf.