November 16, 2017

The Honorable Elaine C. Duke
Acting Secretary of Homeland Security
Department of Homeland Security
Washington, D.C. 20528

Dear Secretary Duke:

We are a group of 54 computer scientists, engineers, mathematicians, and other experts in the use of machine learning, data mining and other advanced techniques for automated decision-making. We write to express our grave concerns regarding Immigration & Customs Enforcement's (ICE) proposed "Extreme Vetting Initiative." Simply put, no computational methods can provide reliable or objective assessments of the traits that ICE seeks to measure. In all likelihood, the proposed system would be inaccurate and biased. We urge you to reconsider this program.

According to its Statement of Objectives,[1] the Extreme Vetting Initiative seeks to make "determinations via automation" about whether an individual will become a "positively contributing member of society" and will "contribute to the national interests." As far as we are aware, neither the federal government nor anyone else has defined, much less attempted to quantify, these characteristics.[2] Algorithms designed to predict these undefined qualities could be used to arbitrarily flag groups of immigrants under a veneer of objectivity.

Inevitably, because these characteristics are difficult (if not impossible) to define and measure, any algorithm will depend on "proxies" that are more easily observed and may bear little or no relationship to the characteristics of interest. For example, developers could stipulate that a Facebook post criticizing U.S. foreign policy would identify a visa applicant as a threat to national interests.[3] They could also treat income as a proxy for a person's contributions to society, despite the fact that financial compensation fails to adequately capture people's roles in their communities or the economy.

---

[1] U.S. Immigration & Customs Enforcement, "Extreme Vetting Initiative: STATEMENT OF OBJECTIVES," June 12, 2017, *available at* https://www.fbo.gov/utils/view?id=533b20bf028d2289633d786dc45822f1.
[2] *See* David A. Martin, "Trump's 'refugee ban' - annotated by a former top Department of Homeland Security lawyer," *Vox*, Jan. 30, 2017 (referring to these requirements as "remarkably vague criteria that will be very hard to turn into operational guidance").
[3] *See* U.S. Immigration & Customs Enforcement, "Background," June 12, 2017, *available at* https://www.fbo.gov/utils/view?id=3a1078ca9739319d84f05424dd05ef6a ("Task 3: Social Media Exploitation").

The Extreme Vetting Initiative also aims to make automated determinations about whether an immigrant "intends to commit" terrorism or other crime. However, there is a wealth of literature demonstrating that even the "best" automated decision-making models generate an unacceptable number of errors when predicting rare events.[4] On the scale of the American population and immigration rates, criminal acts are relatively rare, and terrorist acts are extremely rare.[5] The frequency of individuals' "contribut[ing] to national interests" is unknown. As a result, even the most accurate possible model would generate a very large number of false positives - innocent individuals falsely identified as presenting a risk of crime or terrorism who would face serious repercussions not connected to their real level of risk.

Data mining is a powerful tool. Appropriately harnessed, it can do great good for American industry, medicine, and society. And we recognize that the federal government must enforce immigration laws and maintain national security. But the approach set forth by ICE is neither appropriate nor feasible.

We respectfully urge you to abandon the Extreme Vetting Initiative.

Sincerely,

Hal Abelson, Massachusetts Institute of Technology

Ben Adida, Clever

Blaise Agüera y Arcas, Google / Machine Intelligence

Solon Barocas, Cornell University

Steven M. Bellovin, Columbia University

danah boyd, Microsoft Research / Data & Society

Elizabeth Bradley, University of Colorado, Boulder / Santa Fe Institute

Meredith Broussard, New York University

Emma Brunskill, Stanford University

Carlos Castillo, Universitat Pompeu Fabra

---

[4] *See*, *e.g.*, The MITRE Corporation, JASON Program Office, *Rare Events*, Oct. 2009 ("There is no credible approach that has been documented to date to accurately anticipate the existence and characterization of WMD-T [weapons of mass destruction-terrorism] threats."); National Research Council of the National Academies of Science, *Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment*, 2008 (finding that terrorist identification via data mining or "any other known methodology" was "neither feasible as an objective nor desirable as a goal of technology development efforts").

[5] For example, from 1975 to 2015, the likelihood of an American dying in a terror attack on U.S. soil was 1 in 3.6 million per year. *See* Alex Nowrasteh, *Terrorism and Immigration: A Risk Analysis*, Cato Institute, Sept. 13, 2016.

Aaron Clauset, University of Colorado, Boulder

Lorrie Faith Cranor, Carnegie Mellon University

Kate Crawford, AI Now, New York University / Microsoft Research

Hal Daumé III, University of Maryland / Microsoft Research

Fernando Diaz, Spotify

Peter Eckersley, Electronic Frontier Foundation

Michael Ekstrand, Boise State University

David Evans, University of Virginia

Ed Felten, Princeton University

Sorelle Friedler, Haverford College

Timnit Gebru, Microsoft Research

Joe Hall, Center for Democracy & Technology

Brent Hecht, Northwestern University

James Hendler, Rensselaer Polythechnic University

Subbarao Kambhampati, Association for the Advancement of Artificial Intelligence / Arizona State University

Joshua A. Kroll, University of California at Berkeley

Been Kim, Google Brain

Susan Landau, Tufts University

Kristian Lum, Human Rights Data Analysis Group

Sascha Meinrath, X-Lab / Penn State University

Alan Mislove, Northeastern University

Margaret Mitchell, Google Research / Machine Intelligence

Deirdre Mulligan, University of California at Berkeley

Cristopher Moore, Santa Fe Institute

Ramez Naam, technologist and author, *The Nexus Trilogy*

Cathy O'Neil, mathematician and author, *Weapons of Math Destruction*

Jake Porway, DataKind

Megan Price, Human Rights Data Analysis Group

Gireeja Ranade, Microsoft Research

David Robinson, Upturn

Salvatore Ruggieri, University of Pisa, Italy

Stuart Russell, University of California at Berkeley

Bruce Schneier, Harvard Kennedy School

Cosma Shalizi, Carnegie Mellon University

Julia Stoyanovich, Drexel University

Ashkan Soltani, independent researcher and technologist

Peter Szolovits, Massachusetts Institute of Technology

Hanna Wallach, Microsoft Research / University of Massachusetts Amherst

Nicholas Weaver, International Computer Science Institute / University of California at Berkeley

Meredith Whittaker, AI Now, New York University / Google Open Research

Christo Wilson, Northeastern University

Chris Wiggins, Columbia University

David H. Wolpert, Santa Fe Institute

Rebecca Wright, Rutgers University

** Affiliations for identification purposes only.*