KEEPING THE INTERNET
OPEN • INNOVATIVE • FREE

www.cdt.org

CENTER FOR DEMOCRACY
& TECHNOLOGY

1634 I Street, NW
Suite 1100
Washington, DC 20006

DATE	**1/30/15**
TO	**Transportation Security Administration (TSA)**
FROM	**Center for Democracy & Technology**
# PAGES	

January 30, 2015

Attn: Transportation Security Administration (TSA)
Docket Management Facility
U.S. Department of Transportation
1200 New Jersey Avenue SE.
West Building Ground Floor, Room W12-140
Washington, DC 20590-0001

RE: Comments to the Transportation Security Administration (TSA) on "TSA PreCheck Application Program: Expansion of Enrollment Options," Document Number 2014-30639.

The Center for Democracy & Technology (CDT) is submitting these comments in response to the Transportation Security Administration's (TSA) Request for Information (RFI) on the expansion of enrollment options for the PreCheck program. We believe it will harm personal privacy while doing little or nothing to preserve our national security.

Predicting violence among passengers is a difficult task both because of the small amount of relevant historical data and because this behavior is, almost by definition, irrational. A National Academy of Science Report from 2008 reflects this reality; "automated identification of terrorists through data mining (or any other known methodology) is neither feasible as an objective nor desirable as a goal of technology development efforts."[1] Additionally, the program poses substantial risks for privacy and civil liberties by proposing to include commercially available data. The quality of this data is poor and using it to vet passengers could create significant harms for those passengers applying to the program.

Expanding the use of predictive analytics for basic security may speed the security lines, but it is not a panacea for threat mitigation. Algorithms have great power to infer statistical relationships among huge amounts of data and make predictions based on characteristics of people who match those in the existing data set. This technology can be an effective way to make predictions of future behavior based on past behavior; for example, if you buy a gallon of milk every Monday you will probably buy one next Monday. That is why machine learning is also effective at analyzing billions of data points and looking for patterns such as indicators of credit card fraud.

[1] "Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment." National Academy of Sciences. Page 4. 2008. http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=10072008A

However, “algorithms usually aren't very good at predicting, analyzing, or gaming irrational human behavior."[2] Machine-learning systems cannot innovate and they will never catch the "unknown unknowns" of violent conspiracies—previous attacks such as Pearl Harbor and the attacks of 9/11 have relied on finding flaws in the assumptions of United States security personnel.[3] This hard limitation of current predictive technology is important, primarily as a check on the idea that data is sufficiently predictive of future violent behavior to ensure national security. As security expert Bruce Schneider explains, these types of programs are based on “the dangerous myth that terrorists match a particular profile and that we can somehow pick terrorists out of a crowd if we only can identify everyone.”[4]

In addition to the security concerns, CDT believes that using commercially available data creates privacy and civil liberties harms which can be broken down into four concerns: transparency, accountability, fairness and consent.

1) Transparency: What will the public know about which data is ingested and how it is weighted?

   We recommend providing to the public an explanation for why each stream of data is included in the commercial segment of this program. The explanation must be logical and explained in straightforward language. Additionally, this should include information on how various pieces of data are weighted within your analysis. For example, what degree of power would the profile produced by a data broker have in comparison to an FBI background check? A useful example of the standards for incorporating controls for commercial data can be found in the Office of Management and Budget memorandum on the use of commercial data in the Do Not Pay (DNP) Initiative.[5] As part of this guidance, the Office of Management and Budget (OMB) mandates:

   a. Information in commercial databases must be relevant and necessary to meet the objectives described in section 5 of Improper Payments Elimination and Recovery Improvement Act of 2012 (IPERIA).
   b. Information in commercial databases must be sufficiently accurate, up-to-date, relevant, and complete to ensure fairness to the individual record subjects.
   c. Information in commercial databases must not contain information that describes how any individual exercises rights guaranteed by the First Amendment, unless use of the data is expressly authorized by statute.

   The memorandum also requires that agencies establish rules for accessing the system by agency employees as well as administrative, technical and physical safeguards. Finally, as part of the OMB process on commercial databases, the

---

[2] Steiner, Christopher. “Automate This: How Algorithms Took Over our Markets, Our Jobs, and the World.”

[3] For more on this, see Nate Silver’s “The Signal and the Noise” Chapter 13 “What You Don’t Know Can Hurt You.” April 1, 2013. Penguin Press/Classics.

> *“’You can reasonably predict behavior if people would prefer not to die,’ Rumsfeld told me. ‘But if people are just as happy dying, or feel that it’s a privilege or that it achieves their goal, then they’re going to behave in a very different way.’”*

[4] Schneier, Bruce. “An Easy Path for Terrorists” Boston Globe. August 24, 2004. https://www.schneier.com/essays/archives/2004/08/an_easy_path_for_ter.html

[5] Burwell, Sylvia. “Memorandum for the Heads of Executive Departments and Agencies” Office of Management and Budget. August 16, 2013. http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-20.pdf Page 13.

Treasury Department must provide a written assessment of the suitability of the database[6] – another standard that should be incorporated here. CDT views these controls as the minimum necessary for any system that uses personal information.

2) Fairness: How will you ensure that each data stream and corresponding analysis doesn't have a disparate impact?

Automating decision-making and relying on algorithms does not inherently increase fairness. In fact researchers are discovering that at times the opposite is true. A truism of big data is that an increase in the amount of data will yield more accurate results. But this also creates an important corollary: less information will result in less accurate results. The result of that reality is that big data analysis may never be as accurate for minority populations as it is for majority populations. "It's true by definition that there is always proportionately less data available about minorities. This means that our models about minorities generally tend to be worse than those about the general population."[7] For example, when Google used an algorithm to police its real name policy, several Native Americans had their Google Plus accounts suspended because their names were identified as likely to be fake.[8] In the context of national security, a mistake like this would infringe upon fundamental democratic values.

3) Accountability: What system will be in place to fix mistakes and allow users to audit their own information?

Commercial databases do not face the same incentives as national security professionals with respect to quality control: in the financial system and other avenues where the use of commercial data is routine, we have observed considerable tolerance for error. "The data aggregators are subject to no rules regarding data quality, and their databases are rife with errors, as are the credit ratings agencies' (despite their being subject to some regulations)."[9] However, national security operates with an expectation of zero errors—a missed terrorism signal would weaken security while a falsely identified innocent individual could be subjected to significant harm by being labeled a terrorist. The high stakes nature of these determinations means incorporating this flawed data raises fundamental concerns. Nor does the proposal seem to address other basic due process consideration. For example, in the event that a passenger fails to qualify, is there some way for individuals to see and correct any misinformation? Additionally what happens when an individual fails to pass screening? Does that failure have negative consequences? Could they face additional screening or placement on a watch list?

---

[6] Burwell, Sylvia. "Memorandum for the Heads of Executive Departments and Agencies" Office of Management and Budget. August 16, 2013. Page 7.

[7] Hardt, Moritz. "How big data is unfair: Understanding sources of unfairness in data driven decision making." 9/26/2014. https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de

[8] Flood, Joe. "What Happens When Google Doesn't Think You're A Human" Buzzfeed 3/6/2014. http://www.buzzfeed.com/joeflood/what-happens-when-google-doesnt-think-youre-a-human#.kxBMbgknE

[9] Stanley, Jay. "TSA Once Again Considering Using Commercial Data To Profile Passengers." ACLU Blog. 1/11/2013 https://www.aclu.org/blog/national-security-technology-and-liberty/tsa-once-again-considering-using-commercial-data

4) Consent: Can individuals provide truly informed consent?

One of the factors that has always mitigated the privacy impact of the PreCheck program is its voluntary nature. Privacy risks are substantially lessened when someone willingly enters into an agreement to share their data. However, the significant nature of these very real and unanswered privacy questions described above undermines this factor. Put plainly, can an individual truly consent to participation in a program with so many unknowns when they don't know how their data will be analyzed and what the consequences of their participation might be?

Given these significant and unresolved questions, as well as the program's questionable efficacy, we recommend that TSA not move forward with an expanded PreCheck program. If you have any follow-up questions, please feel free to contact us at 202.637.9800.

Chris Calabrese
Senior Policy Director

Alethea Lange
Policy Analyst
Consumer Privacy Project