Improvements, Questions, and Limits for the Future of Watchlisting

By Timothy Goyer

THE ISSUE
Secretary of Defense Mark Esper promised reforms to the foreign military student program following his visit to Pensacola Naval Air Station in January. The December attack at the Naval Air Station by a Saudi air force student showed the finite ability of the government to continuously vet and assess the risk of foreign visitors—even those sponsored by their own government and invited to U.S. military facilities. President Donald Trump has issued two National Security Presidential Memoranda on vetting, expanding its breadth and giving form to campaign promises of “extreme vetting,” but limits on the predictive capacity of vetting and increasing varieties and volume of data combine to make vetting increasingly difficult.

THE TERRORIST WATCHLIST
More than one million names comprise the Terrorist Screening Database, encompassing persons precluded from flying, those subject to additional screening, and those merely included on the list. The watchlist is integrated with visa, border, and other screening functions and relies upon input from intelligence, law enforcement, and screening agencies.

INTRODUCTION
Vetting of foreign visitors has become a key element of counterterrorism efforts since the attacks of September 11, 2001, the attempted attack on Northwest 253 on December 25, 2009 (the would-be underwear bomber), and through to the potential release of foreign fighters of the Islamic State. The quickening pace of new kinds of data available to be screened and the challenge posed by visitors already present in the United States make the Pensacola attack a cautionary tale. The expansion of vetting, directed in both National Security Presidential Memorandum (NSPM) 7 and 9, is an opportunity to consider the efficiency and effectiveness of the current process and what will need to change as data, computing power, and screening expand. This is in addition to the long-standing Homeland Security Presidential Directive-6, which has served as the originating guidance for the government since 2003.

WHO IS KEPT OUT—AND HOW?
The attack at Naval Air Station Pensacola, the potential release of detained Islamic State and foreign fighters in Syria, a recent court decision in Virginia, and two NSPMs have renewed focus on the terrorist watchlist and its effectiveness in identifying known bad actors. When President Trump campaigned for office, he promised “extreme vetting” of suspected bad actors. And as president, he has issued NSPMs to implement additional kinds of screening and to vet for threats by previously unknown persons. The failure to prevent the entry of suspected terrorists before 9/11 led to the creation of a common watchlist of “known or suspected terrorists.” That
“watchlisting process” was re-examined following the attempted bombing of Northwest Flight 253 on December 25, 2009.

The guidance for the watching enterprise stems from three primary documents:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Issue Date: September 26, 2003.</td>
<td>• Issue Date: October 4, 2017.</td>
<td>• Issue Date: February 6, 2018.</td>
</tr>
<tr>
<td>• Purpose: To enhance the ability of the United States to “develop, integrate, and maintain thorough, accurate and current information about individuals . . . related to terrorism.”</td>
<td>• Purpose: Outlines five new threat categories (weapons proliferators, transnational criminals, foreign intelligence, military, and cyber threat actors) for additional vetting.</td>
<td>• Purpose: Establishes a “National Vetting Center” to “coordinate agency vetting to identify individuals who present a threat to national security, border security, homeland security, or public safety.”</td>
</tr>
<tr>
<td>• The Interagency Terrorist Screening Center functions as the operational arm of the terrorist watchlist.</td>
<td>• The assistant to the president for Homeland Security and Counterterrorism oversees “policy aspects and use of national security threat actor information.”</td>
<td>• This effort is overseen by a National Vetting Governance Board to coordinate the management and governance of the national vetting enterprise.</td>
</tr>
</tbody>
</table>

Homeland Security Presidential Decision-6 (HSPD-6), issued in 2003, mandated extensive terrorism screening and serves as the framework for the existing screening program. Generally, NSPM-7 and NSPM-9 directly work on systems and processes to strengthen the use of intelligence and law enforcement information for national, border, and homeland security purposes. NSPM-7, issued in October 2017, outlines five new threat actor categories (weapons proliferators, transnational criminals, foreign intelligence, military, and cyber threat actors) for additional vetting. NSPM-9, issued in February 2018, establishes a National Vetting Center.

As vetting has expanded, we should examine the effectiveness of the effort thus far. The intersection of the watchlist with many computing developments—from biometric capabilities (e.g., facial recognition, and DNA) to aggregation of open-source information and social media—make this a critical policy question that affects travelers, immigration beneficiaries, and many others. We should pose questions about risk tolerance, efficacy, and goals—questions that like much of the broad-based counterterrorism mission often treated only in absolutes.

What can be learned from implementation of HSPD-6 over the past 15 years as these NSPMs come into force? Are we best to abide likely legal changes (the ruling in September called for additional filings from both the government and plaintiffs) in a system largely exempt from disclosure under the Freedom of Information Act?

It is instructive to understand the current system to identify known or suspected terrorists as they seek visas, entry to the United States, or other benefits, such as lawful immigration status. Through a multi-step process, intelligence and law enforcement agencies provide information about known and suspected terrorists to screening and other agencies. Information gleaned from a wide array of sources is made part of a larger system for screening that has extensive impact upon travelers, applicants for immigration, and others each day.

**HOW MANY TO VET?**

The number of lawful travelers to the United States is vast. More than a million people arrive each day at more than 300 ports-of-entry. Add to this the number of visa exempt applications each year—more than 21 million—plus another 9 million visa applications, as well as requests for permanent residence (low millions), and the scale of vetting becomes clear. This represents only travel and immigration, excluding state and local police forces, the Department of Defense, nuclear facilities, and other watchlist users.

**LEARNING FROM YESTERDAY**

In part, the singular, present-day terrorist watchlist was developed in response to the many independent “terrorist
watchlists” housed within the U.S. government—each with different standards, and applications—that existed prior to the attacks of September 11. Since its inception, the list has expanded, both in the sheer number of names and applicable uses. While policies guiding inclusion of information have become more detailed, the standard for evaluating “nominations” to the watchlist is relatively low (reasonable suspicion) as compared with predication of criminal charges (probable cause). However, the watchlist is not a “wanted” list for persons who meet the standard for an arrest warrant. Rather, it charts a middle path between those without any known activity of concern and those subject to arrest. It is meant to ensure that the reasonable suspicion standard is considered in travel and other kinds of government adjudication. The unstated corollary is that this should be accomplished without needlessly denying applicants or questioning people. This has become steadily more difficult due to a sustained zero-tolerance approach to risk, the vastly increasing available data (e.g., social media, open-source, and biometric), and a lack of assessment of the utility of the current system.

Containing more than a million names, the watchlist simultaneously serves as a screening mechanism, analytic resource, and, implicitly, a deterrent for known or suspected terrorists. It has been refined to better address the problems that any name-based screening system might encounter. For example, name variants, aliases, lack of a common standard for transliteration, lack of biometrics, and thin derogatory information combine to challenge reliable identification on an industrial scale. A formalized redress system has been implemented, but there remains plenty of ways to improve matching, reducing the number of false positives (matching a person incorrectly to the watchlist) and false negatives (missing the person you wanted to match), while maintaining a defensible—legally and practically—mechanism. While the size of the list does not directly correlate to encounters by travelers and others, it is reasonable to infer that the scale contributes to the hurdles of travel and makes assessment of its impact even more vital.

New kinds of data, such as information gleaned from social media, which was scarcely conceived when HSPD-6 was written in 2003, could easily swamp the system. Even prior to the emergence of social media, effectively using open-source information had been a recurring problem. The lack of rigor in considering the impact of such data in the terrorist watchlist today, while adding five new lists, should give pause.

**LEARNING FROM TODAY**

While there is much to learn from the terrorism watchlist as the NSPMs are implemented, the kinds of data used to populate it, the data science necessary to effectively use it, and data overload issues that are common problems throughout the intelligence community also apply here. These represent the mechanical and arguably easier questions. More difficult are questions of utility, cost, and implementation of potential legal decisions, for example. Add to this the glacial pace of government information technology development—be it to store and access data, correlate it accurately, export attributes, or simply have plain but powerful user interfaces—and progress becomes ponderous. Given the request for input by the court, now is a good time to devise changes to solidify the mechanical, practical, and procedural practices.

Despite the size of the list, its impact—both false positives (identifying the wrong person) can be diminished and the fear of false negatives (missing the person you tried to identify) likewise diminished. The zero-tolerance approach needs a policy and operational re-examination. Both kinds of falsities incur costs on agencies and individuals to ensure that no potential persons of interest are missed due to spelling, transliteration, or other mistakes. International travelers account for more than a million screening occasions each day. It is inevitable that mistakes will be made. However, the level of effort in place now exceeds the reasonableness test and is increasingly expensive. Too little
attention is paid to this excess, also without accounting for the impact upon travelers, visa immigration, and other applicants. Additional rigor in assembling the list, greater precision in matching, and more accurate quantification of costs are all achievable goals that maintain compelling government interests and reduce unnecessary screening. More importantly, the process should reduce impact on travelers and applicants.

A RECOMMENDATION AND QUESTIONS FOR EFFECTIVE IMPLEMENTATION OF NSPM-7 AND NSPM-9

Consider a cost-benefit analysis to better guide the screening agencies; an assessment is needed that evaluates the costs associated with additional demands on the screening system.

Questions to begin this examination:

• How many false positives do we accept under the premise that such errors somehow reduce the dreaded false negative?

• How do we measure whether the mere existence of the system acts as a deterrent to bad actors—a laudable outcome—or simply dissuades otherwise eligible persons from even trying to come to the United States?

• How much additional data on persons on the watchlist translates to correct matches? This consideration needs to include new kinds of information, such as social media. However, in stating such, it is important to consider its real import on inclusion and usefulness as an indicator of potential risk. For instance, a one-time casual contact (which would need definition) with a bad actor may not make an individual a bad person, but does that calculation change with repeated, one-on-one communication?

• How much effort is worth the cost of translating social media before it can be correlated, examined for sentiment, or otherwise analyzed? Even social media not needing translation can be exceedingly difficult to assess.

We should think carefully before sweeping vast quantities of interactions onto the files of a person for any watchlist, but we also cannot turn a blind eye to the potential of such data. It is easy to do more, but it is much harder, once begun, to reduce unnecessary screening and harder still to quantify the risk equation when you only add data to the process and presume risk is reduced by information availability. The Governance Board model, outlined in NSPM-9, lacks a strategic capacity. It may not be the best way to fully consider the impact and effectiveness of either the existing process or which efforts are likely to increase the amount of data used and the types of populations that are vetted.

• Measure effectiveness against new kinds of threat actors. We need to analyze whether “watchlisting” will work to any degree with cyber actors. Do cyber actors try to come to the United States in person or is the NSPM contemplating a watchlist that works—somehow—in cyberspace? Even more than the many problems in accurately identifying known or suspected terrorists, correctly matching cyber identities to the watchlist or cyber-to-physical identities would seem to present a very complicated challenge. The operational deconfliction between law enforcement and intelligence on terrorism cases, handled at the Terrorist Screening Center, will be another crucial area to resolve before flagging persons. The cyber sphere will pose issues not evident in the other threat actor categories.

• Plan for legal decisions that will likely expand the review of “nominations” of persons to the watchlist to ensure that “reasonable suspicion” (the current standard for inclusion on the terrorist watchlist) is met but also fulfills the commonsense test (i.e., ensuring that persons reasonably suspected of posing a threat, in an articulable way, be positively identified by screening agencies). As five new watchlists are created, the evolving standards of the terrorist watchlist should serve as the minimal standards for the others. There remains a “compelling interest” for the government in excluding non-citizens and non-residents from entry who are linked to terror or other threat actor activities while reasonably handling a small number of Americans and permanent residents during travel and immigration processing. While incurring cost of time at the outset of the process, additional review for the small number of Americans and permanent residents (thousands of the million plus) may improve efficiency in the matching process and reduce the false positives and better guide those who are correctly identified.

Review of the pattern and practice of threat actors, with emphasis on their access to the United States and susceptibility to screening as envisioned in the NSPMs, would further enhance the utility of any list.

• Apply capabilities of information technology (IT) as a first consideration for both NSPMs, rather than leaving it to post-implementation development. Storage, disambiguation, interoperability, access, and retrieval of data, along with strict security protocols
for users, must precede structure and sharing of the content. Smart IT planning complements good policy development. Capabilities in this sphere are advancing so quickly—and government uptake is so slow—that planning for integration of future advances must be part of any development. The potential legal outcome, almost certain to involve an additional level of quality assurance and justification for placing Americans on the list, could partially be addressed by better computing and IT. Disambiguating names, whether by including biometrics or other identity attributes that positively identify the subjects intended to be matched (thus reducing false positives), requires improvement to computing power and programs, just as now there is daily quality assurance to make sure records are complete and identify the intended person. Information widely shared, as directed by the NSPM-7, is particularly important for future environment considerations.

It is easy to do more, but it is much harder, once begun, to reduce unnecessary screening and harder still to quantify the risk equation when you only add data to the process and presume risk is reduced by information availability.

**REVIEW, REVISE, AND RECOGNIZE**

While neither NSPM calls for a comprehensive review of watchlist performance and policy, it would well serve the public, especially travelers, to re-examine the effectiveness, efficiency, and threat environment against expectations. There will almost certainly be cases in which retrospective review of recent or long-ago comments on social media or a link to a “bad” person will be uncovered, and we should think now about the measure, even if imperfect, to evaluate the meaning of such connections before we become entirely beholden to simple, binary choices about inclusion on a watchlist, entry to the United States, suitability for immigration, or other benefits. Even as the threat environment changes, we will want to perform due diligence in a way that should at best lightly impact travelers and applicants. We should be able to better combine precision and application.

We need as well to understand that given the scale of the screening system, from intelligence collection on counter-terrorism to law enforcement investigations to border and immigration processing, there will be errors, and hindsight will reveal marginal steps to address gaps. We do not live in a zero-risk world and need to stop imagining that vetting immigration and traveler populations will reduce risk to zero. Vetting can reduce, not eliminate, the threat posed by known bad actors and potentially dissuade some unknown bad actors from trying their luck. Even an attack similar to the murders at Naval Air Station Pensacola remains possible. Adapting to new kinds of information and legal challenges from cases, as in Virginia, is key to successful implementation of both NSPMs. Applying the same techniques to different kinds of bad actors, with rapidly evolving computing capabilities in a measured way that tangibly improves the “vetting enterprise” without simply making it even more ponderous, requires a good deal more effort to realize than we have seen since issuance of the NSPMs.

There is a compelling need to carefully synthesize intelligence and law enforcement information and make it available for screening agencies. As was the intent of HSPD-6, robust, integrated screening was, and remains, a necessary goal after September 11, the attempted attack on Northwest 253, and the threat posed by Islamic State foreign fighters. But the need (based on the threat) can be better quantified, the execution made commensurate with the need, and the impact reduced.

**Timothy Goyer** is a visiting fellow with the International Security Program at the Center for Strategic and International Studies in Washington, D.C.

The author would like to give special thanks to CSIS colleagues Asya Akca, Joseph Federici, and Beverly Kirk for their reviews and suggestions.

This report is made possible by general support to CSIS. No direct sponsorship contributed to this report.
ENDNOTES


12 The Department of Homeland Security has proposed a rule for social media information on travel forms, see F.R. Vol. 84, No. 171, Docket No. DHS-2019-0044.