THE OTHER BLIND SPOT IN NSA'S EO 12333 PRIVACY REPORT: RESEARCH

Yesterday, I laid out the biggest reason the NSA Privacy Officer's report on EO 12333 was useless: she excluded most of NSA's EO 12333 collection — its temporary bulk collection done to feed XKeyscore and its more permanent bulk collection done to hunt terrorists and most other NSA targets — from her report. Instead, Privacy Officer Rebecca Richards' report only covered a very limited part of NSA's EO 12333 spying, that targeting people like Angela Merkel.

But I wanted to circle back and note two other things she did which I find telling.

First, note what Richards didn't do.
The standard by which she measured NSA's privacy efforts is a NIST standard called Fair Information Practice Principles, which include the following:

- Transparency
- Individual Participation
- Purpose Specification
- Data Minimization
- Use Limitation
- Data Quality and Integrity
- Security
- Accountability and Auditing

She dismisses the first two because NSA is a spook organization.

> Because NSA has a national security mission, the principles of *Transparency* and *Individual Participation* are not implemented in the same manner they are in organizations

> with a more public facing mission.

In the process, she overstates how assiduously NSA lets Congress or DOJ review EO 12333 activities.

For the rest, however, Richards doesn't — as she should have — assess NSA's compliance with each category. Had she done so, she would have had to admit that PCLOB found NSA's retention under the Foreign Intelligence purpose to be far too broad, putting NSA in violation of Purpose Specification; she would have had to admit that NSA gets around Use Limitation with broad permissions to create technical databases and keep all encrypted communications; she would have had to admit that of NSA's violations, 9% constitute a willful refusal to follow Standard Operating Procedures, a stat that would seem to belie her Accountability claims.

Rather than assessing whether NSA complies with these principles, then, Richards simply checks them off at the end of each of several sections on the SIGINT Production Cycle.

> ACQUIRE, Targeting: "The existing civil liberties and privacy protections fall into the following FIPPs: Transparency (to overseers), Purpose Specification, and Accountability and Auditing."
>
> ACQUIRE, Collection and Processing: "The existing civil liberties and privacy protections fall into three FIPPs categories: Data Minimization, Purpose Specification and Accounting and Auditing."
>
> ANALYZE: "These existing civil liberties and privacy protections fall into the following FIPPs: Transparency (to overseers), Purpose Specification, Data Minimization, and Accountability and Auditing."
>
> RETAIN: "These existing civil liberties and privacy protections fall into the

> following two FIPPs: Data Minimization, and Security."
>
> DISSEMINATE: "The existing civil liberties and privacy protections fall into the following FIPPs: Use Limitations, Data Minimization, and Accountability and Auditing."

Then, having laid out how the NSA does some things that fall into some of these boxes at each step of the SIGINT process, she concludes,

> CLPO documented NSA's multiple activities that provide civil liberties and privacy protections for six of the eight FIPPs that are underpinned by its management activities, documented compliance program, and investments in people, training, tools, and technology.

Fact check! Even buying her claim that checking the box for some of these things at each step of the process is adequate to assessing whether it fulfills FIPP, note that she hasn't presented any evidence NSA meets NIST's "Data Quality and Integrity" claim (though that may just be sloppiness on her part, a further testament to the worthlessness of this review).

But there's another huge problem with this approach.

By fulfilling her privacy review by checking the boxes for the SIGINT Production Cycle (just for the targeted stuff, remember, not for the bulk of what NSA does), Richards leaves out all the other things the NSA does with the world's data. Most notably, she doesn't consider the privacy impacts of NSA's research — what is called SIGDEV — which NSA and its partners do with live data. Some of the most aggressive programs revealed by Edward Snowden's leaks — especially to support their hacking and infiltration activities — were SIGDEV presentations. Even on FISA programs, SIGDEV is subjected to nowhere near the amount of auditing

that straight analysis is.

And the most significant known privacy breach in recent years involved the apparent co-mingling of 3,000 files worth of raw Section 215 phone dragnet data with Stellar Wind data on a research server. NSA destroyed it all before anyone could figure out what it was doing there, how it got there, or what scope "3,000" files entailed.

> In my **obsessions** with the poor oversight over the phone dragnet techs, I have pointed to **this description** several times.
>
> > As of 16 February 2012, NSA determined that approximately 3,032 files containing call detail records potentially collected pursuant to prior BR Orders were retained on a server and been collected more than five years ago in violation of the 5-year retention period established for BR collection. Specifically, these files were retained on a server used by technical personnel working with the Business Records metadata to maintain documentation of provider feed data formats and performed background analysis to document why certain contact chaining rules were created. In addition to the BR work, this server also contains information related to the STELLARWIND program and files which do not appear to be related to either of these programs. NSA bases its determination that these files may be in violation of BR 11-191 because of the type of information contained in the files (i.e., call detail records), the access to the

> > server by technical personnel who worked with the BR metadata, and the listed "creation date" for the files. It is possible that these files contain STELLARWIND data, despite the creation date. The STELLARWIND data could have been copied to this server, and that process could have changed the creation date to a timeframe that appears to indicate that they may contain BR metadata.

The NSA just finds raw data mingling with data from the President's illegal program. And that's all the explanation we get for why!

Well, PCLOB **provides** more explanation for why we don't know what happened with that data.

> > In one incident, NSA technical personnel discovered a technical server with nearly 3,000 files containing call detail records that were more than five years old, but that had not been destroyed in accordance with the applicable retention rules. These files were among those used in connection with a migration of call detail records to a new system. Because a single file may contain more than one call detail record, and because the files were promptly destroyed by agency technical personnel, the NSA could not provide an estimate regarding the volume of calling records that were retained beyond the five-year limit. The technical server in question was not available to intelligence analysts.

> This is actually PCLOB being more solicitous in other parts of the report. After all, it's not just that there was a 5 year data retention limit on this data, there was also a mandate that techs destroy data once they're done fiddling with it. So this is a double violation.
>
> And yet NSA's response to finding raw data sitting around places is to destroy it, making it all the more difficult to understand what went on with it?

Richards may be referring to this kind of oopsie when she talks about "spillage" being a risk related to retention.

> The civil liberties and privacy risks related to retention are that NSA (1) may possibly retain data that it is no longer authorized to retain; (2) may possibly fail to completely remove data the Agency was not authorized to acquire; and (3) may potentially lose data because of "spillage," improper intentional disclosure, or malicious exfiltration.

But nowhere does she consider the privacy implications of having a "technical database" data retention exemption even for Section 702 data, and then subjecting that raw data to the most exotic projects NSA's research staff can think of.

And given that she elsewhere relies on President Obama's PPD-28 as if it did anything to protect privacy, note that that policy specifically exempts SIGDEV from its limits.

> Unless otherwise specified, this directive shall apply to signals intelligence activities conducted in order to collect communications or information about communications, except that it shall not apply to signals

> intelligence activities undertaken to test or develop signals intelligence capabilities.

We know NSA doesn't abide by privacy rules for its research function. Not only does that mean a lot of probably legitimate research evades scrutiny, it also creates a space where NSA can conduct spying, in the name of research, that wouldn't fulfill any of these privacy protections.

That's a glaring privacy risk. One she chooses not to mention at all in her report.