THE 702 COMPLIANCE REPORTING

This will be a very weedy post on two quarterly reports on 702 compliance released to ACLU under FOIA: March 2014, March 2015; the March reports both cover the December 1 through February 28 period. ACLU obtained them not by FOIAing quarterly compliance reporting directly. Rather, ACLU asked for all the documents referred in this Summary of Notable Section 702 Requirements, which they had received earlier. But the released copies are entirely useless in elucidating the Notable Requirements. The 2015 report, for example, was provided in part to explain how NSA assesses whether a selector will provide foreign intelligence information, but the section of the report that details with it (item 28 on page 46) has been withheld entirely (see break between PDF 8 and 9). In addition, there must be at least one more citation to it that is redacted in the Notable Requirements document. The reference(s) to the 2014 report are entirely redacted.

There are a few places such redacted references to the two reports might be: There's a missing citation in Pre- and Post-Tasking Due Diligence (the redaction at the bottom of 2). There may be a citation missing in the continued assessment section at the bottom of page 4. There's definitely one missing in the Obligation to Review section (page 5). There's likely to be one in the long redacted passage on page 6 pertaining to resolving post-tasking problems as quickly as possible. And the sole footnote (see page 11) in the Summary has a reference, which is likely one on FBI techniques to analyze Section 702 information the government identified as being withheld in its entirety.

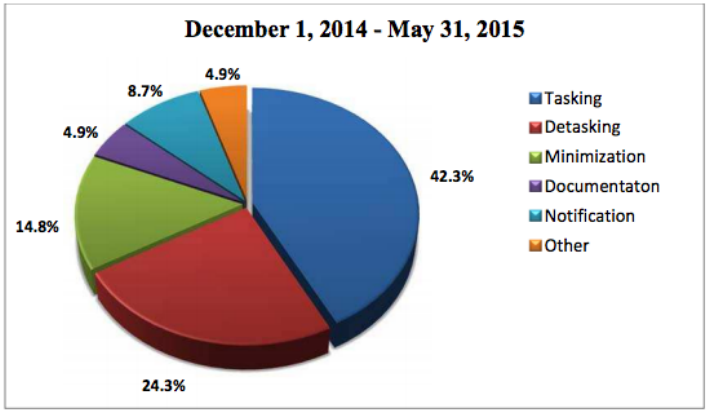
So the Compliance reports don't help us — at all — to understand the requirements the government places on itself with respect to 702.

But they do show us, in more granular detail

than show up in the Semiannual reports (this one includes the March 2014 period and this one includes the March 2015 period), the kinds of things that show up in the compliance reviews. The compliance reporting in both is generally organized in to the same sections (see page 29):

- Tasking Issues
- Detasking Issues
- Notification Delays
- Documentation Issues
- Overcollection
- Minimization
- Other

And — as the Semiannual Report makes clear — we're just seeing a fraction of the granular descriptions in the quarterly reports, because we're not seeing the tasking, detasking, notification, or documentation issues. That means the unredacted content in the released reports represents less than 20% of the total number of compliance incidents for these two quarters.



Though we may be able to use the reports in conjunction to identify how many selectors, on average, are tasked at any given time. If the 25 minimization issues cited in the March 2015 report are representative (meaning there'd be 50 for the entire six month period), then there'd be roughly 338 incidents across all topics for the six month period (it's not entirely clear how they deal with overlap). Given a compliance rate of .35% per average facilities tasked, this

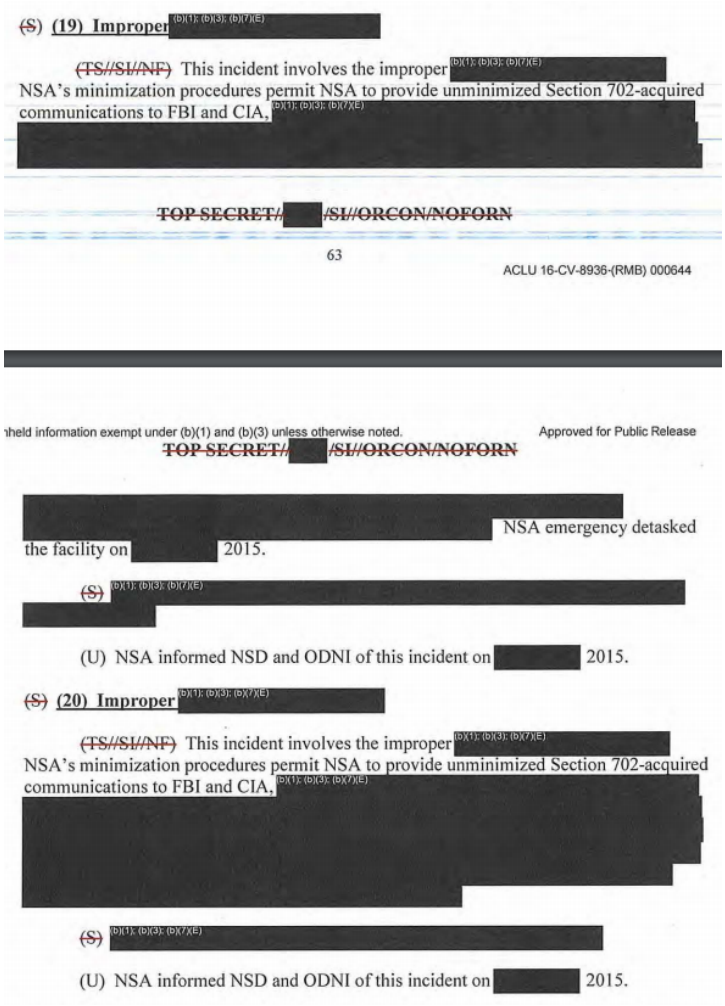
means roughly 96,571 facilities tasked at any given time, thought that may be low given the vastly different lead times on these reports (meaning in the interim year, the government might ID many more compliance issues that get reported primarily in the Semiannual report). There were 94,368 targets across the whole year in FY 2015 (which covers this entire period because the Fiscal Year begins in October). What that suggests is that for some targets, you'll have more than one facility tasked at any given time, but unless there's a lot of turnover in a given year (meaning that most targets are only tasked for some weeks or months), not that many.

Which leaves us with what the reports do show us: the other (largely dissemination) and minimization (largely overly broad queries and US person queries) compliance errors, errors which I've roughly tallied in this document.

Dissemination

Between the two quarterly reports, there are 13 incidences of what I'm lumping under improper dissemination (the report treats database dissemination differently from disseminating unmasked USP identities). Most of these are fairly non-descript, true error. In three cases, analysts at *other* agencies alerted the NSA that they had not masked a US person identity.

The exceptions are 2015-19 and -20, which are almost entirely redacted but pretty clearly deal with NSA sharing raw data with FBI and/or CIA improperly.



~~(S)~~ **(19) Improper** [redacted]

~~(TS//SI//NF)~~ This incident involves the improper [redacted] NSA's minimization procedures permit NSA to provide unminimized Section 702-acquired communications to FBI and CIA, [redacted]

~~TOP SECRET//~~[redacted]~~/SI//ORCON/NOFORN~~

63

ACLU 16-CV-8936-(RMB) 000644

[redacted]

[redacted]held information exempt under (b)(1) and (b)(3) unless otherwise noted. Approved for Public Release

~~TOP SECRET//~~[redacted]~~/SI//ORCON/NOFORN~~

[redacted] NSA emergency detasked the facility on [redacted] 2015.

~~(S)~~ [redacted]

(U) NSA informed NSD and ODNI of this incident on [redacted] 2015.

~~(S)~~ **(20) Improper** [redacted]

~~(TS//SI//NF)~~ This incident involves the improper [redacted] NSA's minimization procedures permit NSA to provide unminimized Section 702-acquired communications to FBI and CIA, [redacted]

~~(S)~~ [redacted]

(U) NSA informed NSD and ODNI of this incident on [redacted] 2015.

I find the second one — which includes no unredacted discussion of emergency detasking or other mitigation — to be the more alarming of the two. But in general, the possibility that NSA might mistakenly send FBI (especially) the wrong data is troubling because once things get to FBI they get far less direct scrutiny (both in terms of compliance reviews and in terms of auditing) than NSA gets. Sending the collection on an entire selector over to another agency is far more intrusive than sending over one unmasked name (though it's not clear this raw data belonged to a US person). Plus, once things get to FBI they can start having repercussions.

Overbroad Queries

The overbroad queries are interesting not so much because they affect US persons directly (though they do in perhaps two cases), but for what they say about the querying process. Here's

what the 2015 Semiannual Report says about overbroad queries, which it acknowledges is a problem even while attributing the problem to errors in constructing Boolean queries.

> (U) NSA's minimization procedures require queries of Section 702-acquired data to be designed in a manner "reasonably likely to return foreign intelligence information." Approximately 29% of the minimization errors in this reporting period involved non-compliance with this rule regarding queries (54% in the last reporting period).56 As with prior Joint Assessments, this is the cause of most compliance incidents involving NSA's minimization procedures. These types of errors are typically traceable to a typographical or comparable error in the construction for the query. For example, an overbroad query can be caused when an analyst mistakenly inserts an "or" instead of an "and" in constructing a Boolean query, and thereby potentially received overbroad results as a result of the query. No incidents of an analyst purposely running a query for nonforeign intelligence reasons against Section 702-acquired data were identified during the reporting period, nor did any of the overbroad queries identified involve the use of a United States person identifier as a query term.

That generally accords with the most common description of the compliance errors: an analyst constructs a query poorly, recognizes as soon as she gets the results (presumably resulting in far more returned records than expected), someone (the reports as often as not don't tell us who) deletes them, and it gets reported. There are a few incidents where analysts run multiple such queries before discovering the problem — that seems like more of a concern, as fat-fingering a Boolean connector shouldn't

explain it. I'm interested in the errors (2015-7, -8, and -9) where the redaction seems to suggest either some other kind of query or some embarrassment about disclosing that top secret method, Boolean search; it's possible this pertains to XKS searches, which can also involve scripts. One of these overboard queries was done by a linguist (which given the Reality Winner case is interesting). There are also discrepancies about whether the analyst themselves discovered the problem or an auditor, the latter of which happened at least five times (two incidences don't describe who discovered them). Finally, there are interesting differences in the description of the coaching that happens after an issue. Sometimes none is described. Most often, the report describes the analyst getting a talking to. But in a number of cases, "personnel," which might be plural, get coaching. I'm interested in when more than one person would get such coaching.

Finally, consider what it means that most of these violations seem to involved multiple authorities, including 702. That's not at all surprising: you'd want to track a target across all the collection you had on the person. But that also includes upstream 702, which may be part of the problem upstream became such a problem.

US Person Queries

Finally, there are the queries using US person identifiers that, for some reason, were improper under the guidelines first approved in 2011. As I've noted, these have been a consistent problem since at least 2013. The Semiannual Report acknowledges this, or at least the problems with searching upstream 702 data, which was prohibited in the 2011 guidelines.

> (U) Additionally, as noted in prior Joint assessments, the joint oversight team believes NSA should assess modifications to systems used to query raw Section 702-acquired data to require

> analysts to identify when they believe they are using a United States person identifier as a query term. Such an improvement, even if it cannot be adopted universally in all NSA systems, could help prevent compliance instances with respect to the use of United States person query terms.59 NSA plans to test and implement this recommendation during calendar year 2016. The new internal compliance control mechanism being developed for NSA data repositories containing unevaluated and unminimized Section 702 information will require analysts to document whether the query being executed against the database includes a known United States person identifier. Once the query is executed, the details concerning the query will be passed to NSA's auditing system of record for post-query review and potential metrics compilation. As part of the testing, NSA will evaluate the accuracy of reporting this number in future Joint Assessments.60

As you review the violations discovered in 2014 and 2015, remember that (as noted in the 2017 702 authorization), these results were in a period where NSA was just discovering far more pervasive problems with US person searches. As it is, in each quarter here, there were 10 or 11 inappropriate US person searches. In 2014, a number of those (2,5, 8, 17) were searches of 702 data using identifiers associated with US persons already targeted under Title I, 704, or 705(b). Just one (5) of the 2015 violations was approved for individual targeting, and that appears to be one of the earlier violations in the quarter (note it must have occurred in December 2014). That's interesting, because this undated guideline on USP queries of 702 collections says *any* US person approved for individualized targeting or RAS (under the old phone dragnet) could be backdoor searched. It seems likely, then, they changed the policy in

2015 (which is particularly alarming, given that they did so just as NSA was moving towards discovering how bad their upstream searches were. In other words, they seem to have made legal one of the practices that was coming up as a violation.

These violation descriptions are also interesting for the (often redacted) specificity about the kind of selector used, sometimes described as email, telephony (which could include messaging), and in others as "facilities" (which might include cookies or IPs). That's an indication of the range of identifiers under which you can search 702 data, which is in turn (because 702 searches are all supposed to derive from PRISM collection) a testament to the kinds of things that get turned over in PRISM returns.

Of the violations described, just one obviously pertains to the search on an identifier for which the authorization had expired. That's interesting, because searches on expired warrants appeared far more frequently in past reports. Significantly, the IG Report reviewing compliance 704/705(b), which reviewed queries for two months that overlapped with the 2015 report at issue (January and February 2015; the compliance report included December 2014 whereas the IG Report included March 2015), did find persistent problems with expired authorizations, but in EO 12333 data (suggesting FISA queries might have fixed earlier such problems). But the discussion of these problems in Rosemary Collyer's 702 reauthorization opinion shows that for one tool, 85% of 704/705(b) queries conducted from November 2015 through April 2016 — well after the later quarter covered here — were non-compliant. "Many of these non-compliant queries involved use of the same identifiers over different date ranges." NSA was unable to segregate and destroy the improper queries. That's perhaps unsurprising, because as late as April 2017, the NSA was still having difficulties identifying all the queries run against 702 data.

And in spite of the reports, from later 702 reporting that some of the 704/705(b) queries of 702 did not get included in auditing systems, a good number of these violations were not discovered by analysts (as often happened with improper queries) but by auditors, suggesting the violations may have had an impact on US persons.

All that said, there's not all that much there there, aside from the sheer number (which the Semiannual report seems to think is just NSA's serial refusal to fix the problem of default search settings). These two snap-shots of the 702 upstream query problem, capturing 702 collection in the period immediately before it started to blow up, are also an indication of how much ODNI/DOJ's oversight of NSA (which is far more rigorous than the oversight than the same agencies give CIA and especially FBI) was missing.