21% OF THE DATABASE QUERY ERRORS IN NSA REPORT INVOLVED THE ~~PHONE~~ INTERNET DRAGNET DATABASE

Update: as Mindrayge notes, Marina appears in NSA

TOP SECRET//COMINT//NOFORN

(U//FOUO) Figure 6: 1QCY11 Database Query Incidents – Database(s) Involved

U//FOUO

|	
CLOUD/ABR	(8)
DISHFIRE	(16)
FASTSCOPE	(1)
MARINA (SSR)	(24)
OCTAVE	(1)
PINWALE (SSR)	(34)
SIGINT NAVIGATOR	(1)
TRACFIN	(2)
TRANSX	(2)
TUNINGFORK	(4)
UTT	(1)
XKEYSCORE	(24)
Unknown	(1)

Total: 119

U//FOUO

slides as Internet, not phone metadata (and that's how Ambinder refers to it here). There are some oddities, then, but I am changing this post accordingly.

As I noted in this post, the May 3, 2012 audit of NSA's violations falsely suggests "roamer" problems were the cause of an increase in incidents, rather than database query errors, transit collection, and detask problems.

Database query errors are basically when an analyst collects too much data because she doesn't exclude data that should be excluded, she ran a query believing it was appropriate because she had too little information on it, or she ignored standard operating procedures.

In addition to telling us how many database query problems there were, the report tells us which NSA databases they involved. As the figure above notes, 24 of those errors involved the MARINA database. There were actually 115 total query errors — 4 involved multiple databases — which means 21% of the database query errors involve MARINA.

~~As Marc Ambinder and others~~ have reported~~,~~ ~~MARINA is the name of the Section 215 phone~~
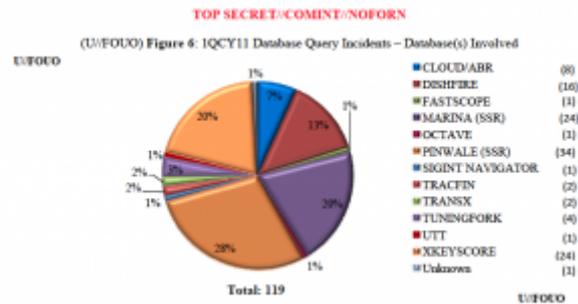
~~records dragnet database.~~

> ~~The telephone metadata is stored in a database called MARINA, which keeps these records for at least five years.~~

In other words, a fifth of the database query errors in the first quarter of 2012 were on the US ~~phone~~ Internet record dragnet database — the one the government has been claiming is so carefully guarded.

[If Mainway is just Internet metadata, then we don't know the number of queries.]

~~Not only that, but we have a rough idea of how common query errors on this database are. The government has told us that queries were made on fewer than 300 identifiers in 2012. While it's not a one-to-one comparison (some identifiers would have been run more than once), that means perhaps as many as 8% of the queries on the dragnet database involved some kind of error, including errors like not following procedures. And that's assuming analysts didn't keep making errors with the database at the same rate they did in the first quarter: if they kept up the same error pace, the error rate might be closer to 32%~~

~~But don't worry, the government tells us, our phone record data are safe, even with a potential error rate of 32% accessing that data.~~

Update: LAT's Ken Dilanian, who listened to a conference call NSA just had, just tweeted this:

> NSA's DeLong will not say how often NSA makes privacy errors when it queries US phone records database. But less than 30%, he says.
>
> I asked is the rate between 8 and 30%, and he said 30% isn't right. So, you may be on to something.

Less than 30%?!?!? That suggests it is probably far higher than even I imagined. Even if it was

8% it would be unacceptably high. But if it's at the higher end of the possible range, it is unbelievably high.

Update: Ron Wyden and Mark Udall have issued a statement on this. Among other statements, they emphasize that Americans need to know about the phone and Internet dragnet violations.

> Americans should know that this confirmation is just the tip of a larger iceberg.
>
> [snip]
>
> In particular, we believe the public deserves to know more about the violations of the secret court orders that have authorized the bulk collection of Americans' phone and email records under the USA PATRIOT Act.

Given the potential numbers of phone dragnet violations, I should say so.

Update: Fixed "a fifth" for "a quarter." Now I'm making NSA type simple math errors!