

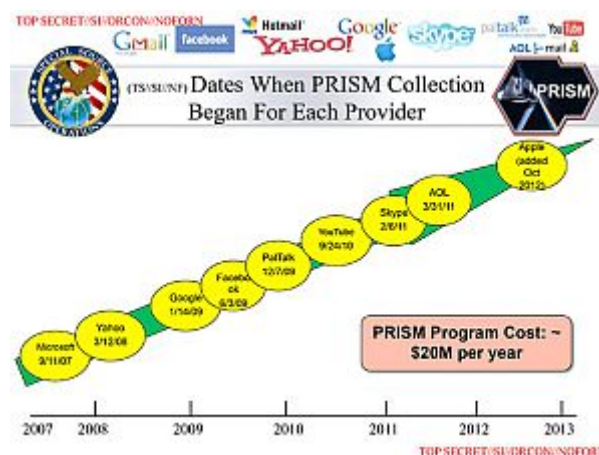
IS GOOGLE SHARING 9,500 USERS' DATA, OR 65,000?

Foreign Intelligence Surveillance Act (FISA) requests		
Non-content requests		
Reporting Period	Number of requests	Users/Accounts
July to December 2013	Data subject to six month reporting delay	
January to June 2013	0-999	0-999
July to December 2012	0-999	0-999
January to June 2012	0-999	0-999
July to December 2011	0-999	0-999
January to June 2011	0-999	0-999
July to December 2010	0-999	0-999
January to June 2010	0-999	0-999
July to December 2009	0-999	0-999
January to June 2009	0-999	0-999

Content requests		
Reporting Period	Number of requests	Users/Accounts
July to December 2013	Data subject to six month reporting delay	
January to June 2013	0-999	9000-9999
July to December 2012	0-999	12000-12999
January to June 2012	0-999	8000-8999
July to December 2011	0-999	9000-9999
January to June 2011	0-999	7000-7999
July to December 2010	0-999	5000-5999
January to June 2010	0-999	3000-3999
July to December 2009	0-999	3000-3999
January to June 2009	0-999	2000-2999

Google just released its shiny new transparency numbers reflecting DOJ's new transparency rules.

While they tell us some interesting things, the numbers show how many questions the transparency system raises. I've raised the questions below, linked to my discussion by bolded number.



Google is using option 1 (perhaps because they had already reported their NSL numbers), in which they break out NSLs separately from FISA orders, but must report in bands of 1000.

Note that Google starts this timeline in 2009, whereas their criminal process numbers pertaining to user accounts only start in 2011. Either because they had these FISA numbers ready at hand, or because they made the effort to go back and get them (whereas they haven't done the same for pre-2011 criminal process numbers), they're giving us more history on their FISA orders than they did on criminal process. They probably did this to show the entire period during which they've been involved in PRISM, which started on January 14, 2009.

Google gets relatively few non-content requests, and the number – which could be zero! – has not risen appreciably since they got involved in PRISM. **(1)** (I suspect we're going to see fairly high non-content requests from Microsoft, because they pushed to break these two categories out).

What Google has gotten are generally increasing numbers of content requests. Some of these are likely to be individual requests. There were 1,788 total FISA applications approved in 2012, of which maybe 8 – 10 are bulk (FAA) requests. So if we assume that Google gets a fair percentage of the individual requests, then maybe 250 of the requests **and users** reflected in Google's content numbers are for individuals. **(2)**

But remember the bulk of what Google turns over is likely under FAA bulk requests (PRISM). We should assume it gets 3-5 requests (one request per certification, of which we know counterterrorism, counterproliferation, and cyber are included), but each might (must) represent thousands of users. So what we're probably looking at are steadily increasing numbers of user accounts affected by NSA/FBI's access of accounts in the name of terror or cyber, backed by just 3-5 requests.

Now look at the pattern, broken only by 2012: The numbers generally rise each year **(3)**, but they spiked in the second half of 2012, then dropped down to earlier levels. Two things might

explain these numbers. First, we know that the aftermath of the upstream collection mess caused Yahoo and Google particular difficulties that took a full year (that is, until the second half of 2012) to clean up. I have speculated that they had to detask a number of identifiers because they were tied to bad MCTs. That might explain both the fall in numbers in the first half of the year, and the big spike at the end of the year. But even as that was happening, NSA had stopped domestic collection of Internet metadata at the end of 2011. It is possible one replacement they used was to get the “metadata” (actually content) via FAA. But the numbers don’t seem to support it—the spike is too late.

Finally, we should at least consider whether these numbers might be cumulative. In the government’s own reporting, they count two numbers at once: the average number of identifiers at any given time, and the number of new identifiers tasked. **(4)** We know the latter number has always risen for Internet content (it fell for phone content in 2009 but is rising again). Thus, it **may be** (we don’t know one way or another and it’s the government’s fault we don’t know) that these numbers only measure the start of someone being wiretapped (again, this is one of the two ways the government counts its own taskings), meaning it’s possible that close to 65,000 people are still being collected under FAA orders. I suspect it’s probably the smaller number, but it is possible that it’s not.

[Update: In my next post, I’ll show that it is probably the smaller number.]

(1) How is “non-content” defined, and will all providers define it in the same way? Did they agree on definitions in some kind of sealed part of the agreement? Because it is conceivable the government will go to Internet providers to get metadata, but much of the metadata they’d want would legally be content. And how would URL searches appear?

(2) What does a FISA request on an individual user entail? If I ask FISA for a warrant for,

say, Anwar al-Awlaki, does one request cover all his communications, across all known platforms? If so, then the 1,788 number may be 1,788 times four or so, including at least one phone platform, plus several Internet providers. Also, we still don't know how 703 and 704 – the warrants for content collection on US persons overseas – work. But they should be in the FISA numbers as well.

(3) We know certificates got approved – at least in 2008 and 2011 – in the fall. Is there a surge connected with that process?

(4) Do the “user/accounts” affected numbers for content reflect identifiers newly tasked, but they can be tasked indefinitely. Or do they affect total users affected at any given time? (The government uses both measures in its own counting).

Update: Here's Yahoo's report, which shows 30-31,000 accounts are affected under FISA requests. This actually makes me think Google's numbers may be cumulative, because I find it hard to believe that NSA would tap three times the number of Yahoo users as Google users, especially given that since Yahoo doesn't default on encryption, it is easier to get their content overseas.

Also note, Yahoo is being misleading here:

The **Number of Accounts** is typically larger than the number of users and accounts involved because an individual user may have multiple accounts that were specified in one or more requests, and if a request specified an account that does not exist, that nonexistent account would nevertheless be included in our count.

It's ignoring that they also would have multiple accounts per request under bulk orders.

Update: Here's Microsoft's. They've had the same 0-999 requests, affecting a total of

15,000-15,999 identifiers.

Particularly helpful is its definitions:

- *FISA Orders Seeking Disclosure of Content:* This category would include any FISA electronic surveillance orders (50 U.S.C. § 1805), FISA search warrants (50 U.S.C. § 1824), and FISA Amendments Act directives (50 U.S.C. §1881) that were received or active during the reporting period.
- *FISA Orders Requesting Disclosure of Non-Content:* This category would include any FISA business records (50 U.S.C. § 1861), commonly referred to as 215 orders, and FISA pen register and trap and trace orders (50 U.S.C. § 1842) that were received or active during the reporting period.
- *Accounts Impacted:* The number of user accounts impacted by FISA orders that were received or

active during the period of time. Since individuals may have multiple accounts across different Microsoft services – all of which are counted separately to determine the number of accounts impacted – this number will likely overstate the number of individuals subject to government orders.

Note it makes clear that it will get both a stored communications search (for archived emails) and traditional content warrants. I can't tell for sure but it seems to suggest that would count as two requests (but we don't know how DOJ reports it).

Finally, note that MS says it has challenged requests.

It is important to remember that receipt of an order does not mean the information that was sought was ultimately disclosed. Microsoft has successfully challenged requests in court, and we will continue to contest orders that we believe lack legal validity.

We've been told that no one has ever challenged a Section 215 order (but who knows if that's true). We know Google has challenged NSLs (though often unsuccessfully). And it's unlikely anyone would challenge a FAA directive itself given Yahoo's failure doing so. So is MS just saying they've challenged NSLs?