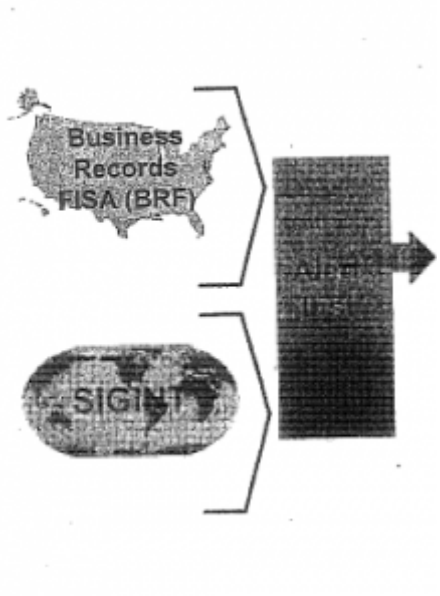
NSA'S LATEST CLAIM: IT ONLY GETS 30% OF "SUBSTANTIALLY ALL" THE HAY IN THE HAYSTACK

> In December 2007, the FBI began intercepting MOALIN's cell phone.
>
> 
>
>
> — FBI search warrant affidavit seeking (among other things) additional cell phones, October 29, 2010

Yesterday, Siobhan Gorman reported that NSA's "phone-data program" collects 20% or less of the phone data in the US. She explains that the program doesn't collect cell phone data, and so has covered a decreasing percentage of US calls over the last several years.

> The National Security Agency's phone-data program, which has been at the center of controversy over the NSA's surveillance operations, collects information from about 20% or less of all U.S. calls—much less than previously described by lawmakers.
>
> The program had been described as collecting records on virtually every

> phone call placed in the U.S., but in fact, it doesn't cover records for most cellphones, the fastest-growing sector in telephony and an area where the agency has struggled to keep pace, according to several people familiar with the program.

Ellen Nakashima's report places the percentage between 20 and 30%, echoing Gorman's claim about limits on cell data.

> The actual percentage of records gathered is somewhere between 20 and 30 percent and reflects Americans' increasing turn away from the use of land lines to cellphones. Officials also have faced technical challenges in preparing the NSA database to handle large amounts of new records without taking in data such as cell tower locations that are not authorized for collection.
>
> [snip]
>
> The bulk collection began largely as a land-line program, focusing on carriers such as AT&T and Verizon Business Network Services. At least two large wireless companies are not covered — Verizon Wireless and T-Mobile U.S., which was first reported by the Wall Street Journal.
>
> Industry officials have speculated that partial foreign ownership has made the NSA reluctant to issue orders to those carriers. But U.S. officials said that was not a reason.
>
> "They're doing business in the United States; they're required to comply with U.S. law," said one senior U.S. official. "A court order is a court order."
>
> Rather, the official said, the drop in

> collection stems from several factors.
>
> Apart from the decline in land-line use, the agency has struggled to prepare its database to handle vast amounts of cellphone data, current and former officials say. For instance, cellphone records may contain geolocation data, which the NSA is not permitted to receive.

These reports offer a more credible explanation than Geoffrey Stone's multiple claims to this effect about why the program misses data. So they may be true.

But I think they instead point to the legal range of authorities NSA uses to collect phone records, not to what records they actually have in their possession.

These reports are commenting (though without specifying, or even seeming to be aware they need to specify) on what the government claims it collects under Section 215. These reports are not commenting on what NSA collects under all authorities.

In this post I will show why I believe these reports to be credible only in a very narrow sense. In a follow-up post I will point to the legal issues that underlie the Administration's conflicting claims about what it collects.

One reason I have always questioned the claim that the government only collects a fraction of US call records is because of past Administration claims about the intent of the program. Nakashima even notes this: back in July James Cole said you have to have the entire haystack to find a needle. Cole is not the only who has made the claim in official settings. Even given James Clapper's history of lying to Congress, I still take comments to Congress with greater weight than anonymous, obviously seeded leaks to reporters at convenient times (to Nakashima's credit, she gets NSA Deputy Director Rick Ledgett on the record, though tellingly, in

his comment, he refuses to say anything about the scope of the program).

The Intelligence Community was implying the collection is comprehensive even before Snowden's leaks. In 2011, the government told Congress it collected "substantially all" of the records from the providers included in Section 215 orders. Claire Eagan repeated that claim in her July 2013 opinion on the dragnet.

> Specifically, the Report provided the following information: 1) the Section 215 production is a program "authorized to collect in bulk certain dialing, routing, addressing and signaling information about telephone calls … but not the content of the calls…." Ex. 3, Report at 1 (emphasis in original); 2) this Court's "orders generally require production of the business records (as described above) relating to **substantially all** of the telephone calls handled by the companies, including both calls made between the United States and a foreign country and calls made entirely within the United States,"

And while Eagan doesn't endorse that claim in her own voice, she raises the "substantially all" language again in her legislative reenactment discussion. She may not be saying NSA collects substantially all phone data, but she is arguing (incorrectly) that Congress authorized it to do so.

While that certainly leaves the possibility of treating Verizon land line service separately from Verizon wireless (it's not clear the Congressional notice ever got that much detail on which providers were included, and there are reasons the Administration may have wanted to claim it got everything from the providers even if it only got land line traffic), the notice to Congress in 2011, repeated last year, supports a claim of much broader collection.

The traditional Article III judges who have reviewed the program similarly seem to believe all the phone data is collected. William Pauley (who presided over a suit naming Verizon Business Services directly) wrote,

> This blunt tool only works because it collects everything.

Richard Leon (who presided over a suit naming Verizon Wireless) captures the conflict within the government's message better.

> The Government obviously wants me to infer that the NSA may not have collected records from Verizon Wireless (or perhaps any other non-VBNS entity, such as AT&T and Sprint). Curiously, the Government makes this argument at the same time it is describing in its pleadings a bulk metadata collection program that can function *only* because it "creates an historical respository that permits restrospective analysis of terrorist-related communications *across multiple telecommunications networks*. and that can be immediately accessed as new terrorist-associated telephone identifiers come to light." Govt.'s Opp'n at 12 (emphasis added); *see also id*. at 65 (removing plaintiff's phone numbers "could undermine the results of any authorized query of a phone number that based on RAS is associated with one of the identified foreign terrorist organizations by eliminating, or cutting off potential call chains").
>
> Put simply, the Government wants it both ways. Virtually all of the Government's briefs and arguments to this Court explain how the Government has acted in good faith to create a comprehensive metadata database that serves as a potentially valuable tool in combating terrorism—in which case, the NSA must have collected metadata from Verizon

> Wireless, the single largest wireless carrier in the United States, as well as AT&T and Spring, the second and third-largest carriers. [snip] Yet in one footnote, the Government asks me to find that plaintiffs lack standing based on the theoretical possibility that the NSA has collected a universe of metadata so incomplete that the program could not possibly serve its putative function. Candor of this type defies common sense and does not exactly inspire confidence!

With Leon's take in mind, consider what we know about the NSA's dragnet.

First, the dragnet that NSA now claims doesn't include cell data seems to keep finding cell phones, even as far back as 2007 when it identified Basaaly Moalin's cell phone off what was almost certainly a cell phone used by Aden Ayro (the NSA would later ask the FBI's help to find Ayro's new phone, suggesting he was using burner cells, and Somali warlords certainly don't operate in an environment with well-developed land line infrastructure). The Najibullah Zazi phone NSA contact chained was a cell phone, and my understanding is the Adis Medunjanin phone the dragnet found was also a cell. The NSA boasted of using the dragnet for peace of mind after the Tsarnaev brothers — at least Dzhokhar of which used his cell phone constantly, even during their attack — attacked the Boston Marathon.

The main examples the NSA offers as phone dragnet successes involve cell phone targets.

That doesn't mean land lines weren't involved — the second hop connecting Aden Ayro and Basaaly Moalin (whom I've always suspected was his *hawala*) — could have been a land line. The second hop between the Zazi cell and what I believe was a Medunjanin cell might be a land line. But the targets here all used cells. There's no reason you'd design a dragnet targeting likely immigrants (as all these men

were) without including the cell phones that reflect their potentially more transient lifestyles (each plot featured men who had worked as cabbies or drivers at some point).

We also know the Section 215 data is just one part of redundant database that also includes EO 12333 data (and probably data from GCHQ). By 2009, NSA identified and tracked 3,000 suspect US phone identifiers using non-Section 215 means (and NSA simply kept them in its EO 12333 dragnet after it discovered they hadn't received First Amendment review). NSA trains analysts to use the redundancy of the system, to create EO 12333 results, if possible, even if originally finding queries via Section 215 data, because the former have more permissive dissemination rules.

And in both the case of Najibullah Zazi …

> This detail, available only at the "second hop" and only visible due to the blending of BR FISA and SIGINT data, quickly identified the Medunjanin number as a priority lead for the FBI.

And David Headley…

> Collection against foreign terrorists and telephony metadata analysis were utilized in tandem with FBI law enforcement authorities to establish Headley's foreign ties and them in context with his U.S. based planning efforts.

… The NSA used both Section 215 and EO 12333 data in its queries (indeed, PCLOB confirms what I noted here — that the useful contact chaining on Headley was conducted under other authorities).

> Further investigation, also not involving Section 215, provided insight into the activities of his overseas associates. In addition, Section 215

> records were queried by the NSA, which passed on telephone numbers to the FBI as leads. Those numbers, however, only corroborated data about telephone calls that the FBI obtained independently through other authorities.

With the recognition that the NSA uses Section 215 in conjunction with EO 12333 data (probably including Internet data), check out how NSA SID Director Theresa Shea alternates between talking about the Section 215 phone data and telephony metadata more generally, and discusses how Section 215 complements other authorities.

> 50. Furthermore, the Section 215 metadata program complements information that the NSA collects via other means and is valuable to NSA, in support of the FBI, for the linking of possible terrorist-related telephone communications that occur between communicants based solely inside the U.S.
>
> 51. As a complementary tool to other intelligence authorities, the NSA's access to telephony metadata improves the likelihood of the Government being able to detect terrorist cell contacts within the U.S. With the metadata collected under Section 215 pursuant to FISC orders, the NSA has the information necessary to perform the call chaining that enables NSA intelligence analysts to obtain a much fuller understanding of the target and, as a result, allows the NSA to provide FBI with a more complete picture of possible terrorist-related activity occurring inside the U.S.

When the NSA talks about a "telehphony metadata" database, they're referring to an interface that draws on data from EO12333 and Section 215, at a minimum. It's the larger dragnet — not just the Section 215 subsection of it — that needs to be

and probably is comprehensive.

What the NSA is anonymously leaking to journalists is that Section 215 only gets a fraction of the US phone data. What the NSA **is not saying** (and what their more formal declarations have made clear they're not saying) is that NSA only has access to 30% of US phone data.

All that said, I think NSA may be leaking what I suspect is deceptive information because of problems they've created for themselves with the dragnet, as I'll show in my follow-up post.

Update: Note that the LAT version (h/t PJ Evans) of this story notes PCLOB Chair David Medine has asserted that the phone dragnet collects everything.

> And in written testimony to the House Judiciary Committee this week, David Medine, chairman of the Privacy and Civil Liberties Oversight Board, which received classified briefings on the NSA systems and issued a lengthy report to Obama last month, said the program involved "ongoing collection of virtually all telephone records of every American."
>
> Medine did not respond to a request for comment Friday.

While in context of his testimony, Medine's assertion seems to refer exclusively to Section 215 phone records, it is worth noting that PCLOB — which as I showed above considered closely how 215 interacts with 12333 data — made this assertion.

Update: This WSJ article from June (which some of these stories cite without thinking through what it means) makes it clear that, while NSA doesn't get T-Mobile and Verizon's cell data from them (suggesting it does get AT&T and Sprint's cell data), it does get that data from other sources.

> The National Security Agency's controversial data program, which seeks to stockpile records on all calls made in the U.S., doesn't collect information directly from T-Mobile USA and Verizon Wireless, in part because of their foreign ownership ties, people familiar with the matter said.
>
> The blind spot for U.S. intelligence is relatively small, according to a U.S. official. Officials believe they can still capture information, or metadata, on 99% of U.S. phone traffic because nearly all calls eventually travel over networks owned by U.S. companies that work with the NSA.
>
> [snip]
>
> When a T-Mobile or Verizon Wireless call is made, it often must travel over one of these networks, requiring the carrier to pay the cable owner. The information related to that transaction—such as the phone numbers involved and length of call—is recorded and can then be passed to the NSA through its existing relationships. Additionally, T-Mobile relies on other wireless companies to fill holes in its infrastructure. That shared equipment could allow the government to collect the data.