

Responses to Questions for the Record
Submitted October 29, 2013
Edward W. Felten
Professor of Computer Science and Public Affairs, Princeton University

United States Senate, Committee on the Judiciary
Hearing on
Continued Oversight of the Foreign Intelligence Surveillance Act
October 2, 2013

I thank the Committee for the opportunity to respond to these Questions for the Record.

Senator Klobuchar's Question

I am very interested in your recommendation that the FISC should have greater in-house technological expertise to assess the government's bulk collection and surveillance requests. I'd like to ask you to flesh this out a bit more.

How would you recommend working technology experts into the current FISC process?

Response to Senator Klobuchar's Question

In the current FISC process, the government is the only party that files papers and argues before the Court. The most natural way to add independent technical expertise would be for the expert to assist the Court. The Court might follow the practice of some ordinary District Courts by retaining a Court-appointed expert, or by appointing a special master who has technical expertise.

If the FISC process is changed to add another party empowered to participate in FISC matters, such as a representative of the public or an advocate for civil liberties, then this party could retain technical experts to assist it in its argument. This expert assistance is important in allowing the independent party to do its job, because the government's argument before the FISC is well-supported by technical experts, and technical claims often play an important role in the government's argument.

If the process is indeed changed to add an independent party, it is important for this party to be able to challenge the government's technical claims. In an ordinary court case, this would occur via discovery, including expert reports, depositions, and cross-examination of experts. Although this full process might not be appropriate for FISC matters, it is important to ensure that the independent party is in a position to get the information it needs to evaluate and challenge technical assertions made by the government.

Finally, the sensitivity of information before the FISC will require that technical experts have the necessary security clearances. Some independent experts already have clearances, but there are relatively few such people who are not already working for or with intelligence agencies. Steps should be taken to make sure that clearance requests can be expedited for technical experts whom the FISC or an independent party want to engage.

Senator Franken's Question

(1) Professor Felten, in your written testimony you stated that "metadata is easy to analyze."

(a) Do you think the intelligence community has the technical ability to give a rough estimate of the number of American citizens and permanent residents whose communications metadata has been collected in their surveillance programs?

(b) Do you think that the intelligence community has the technical ability to give a rough estimate of the number of American citizens and permanent residents whose communications content has been collected in their surveillance programs?

Response to Senator Franken's Question

Yes, the government has the ability to give a rough estimate of the number of American citizens and permanent residents whose (a) metadata and (b) content has been collected.

(a) The intelligence community can give a rough estimate of the number of citizens and permanent residence whose communications *metadata* has been collected. There are several reasonable methods for doing this. Each method gives an estimate that is not exact but is of roughly the correct magnitude.

A first method is to determine the number of U.S. phone numbers that appear in collected metadata records, and then use this information to estimate the number of affected persons. U.S. phone numbers are easily distinguished from non-U.S. numbers by examining the country code and/or area code of the number. Once the number of affected phone numbers is known, this can be used to estimate the number of citizens and permanent residents by making two adjustments, the first to account for the possibility of one person using multiple affected phone numbers, and the second to account for the fact that a small percentage of U.S. phone numbers are owned by people who are neither citizens nor permanent residents.

A second method is to determine the number of distinct customers of each mobile phone carrier whose information is captured. On the assumption that few people have mobile accounts with multiple mobile carriers, this could be used to estimate the total number of affected persons, again correcting for the fact that a small percentage of accounts are owned

by people who are neither citizens nor permanent residents.

A third method, which appears to offer good accuracy if news reports are accurate, is simply to assume that every adult citizen or permanent resident has been on at least one end of a call whose metadata was captured, and therefore to use an estimate equal to the number of adult citizens plus permanent residents.

(b) It is a bit more challenging, but still feasible, for the intelligence community to give a rough estimate of the number of citizens and permanent residents whose communications *content* has been collected.

It is very likely that in all or almost all cases where call content is collected, the metadata about that same call is also collected. If so, then all that remains is to assemble a database of metadata for calls whose content has been captured, and then to use this metadata to estimate the number of affected citizens and permanent residents. This could be done, for example, by using the first method described above in part (a).

Even if, for some reason, content collection is not accompanied by metadata collection for the same calls, it would be feasible to estimate the number of affected citizens and U.S. persons, using the existing metadata.

This is not meant as an exhaustive list of methods, and there are probably better and more accurate methods than the ones I have described here. The intelligence community employs a great many mathematicians, statisticians, and computer scientists, and prides itself on its ability to extract useful information from large data sets. Surely they are able to provide at least rough estimates of how many Americans are affected by their data collection.