

Statement of Barry Steinhardt

1/25/2012

INTRODUCTION

My thanks to the Committee leadership – Richard Purcell and Lisa Sotto -- and the DHS staff for allowing me to submit this separate statement explaining my vote and my concerns regarding the report and DPIAC's limited charge.

This is a profoundly important issue. DHS is a uniquely far flung entity within the Federal Government. A systematic search of its records combined with other governmental data and the commercial information at its disposal would create an unprecedented system of data surveillance.

That said, although I voted "No" on the motion to send the majority report to DHS, this statement is more in the way of a partial concurrence than a pure dissent.

I agree with much of the Majority's document. But, as explained below, I could not vote in favor of a report that contained conflicting statements regarding the critical question of the creation of a new centralized database containing search results.

A. MY VOTE

I am a member of the Policy Subcommittee. While I have reservations about the limitations of the report (see section B below), I supported the Subcommittee's report, which was incorporated into the full Committee's final paper as the " Privacy Policy Guidance".

I also supported the Committee's overarching decision to combine the two Subcommittee reports into one paper with an explanatory preamble.

However, the report of the Technical Subcommittee contained two fundamental assumptions/conclusions which were at odds with those reached by our subcommittee and which went to the heart of the issues raised by the creation of federated information system:

1. Whether queries/searches could be initiated based on generalized information about the nature of the "persons of interest" or whether they would be restricted to queries regarding specific individuals based on personally identifiable information (PII), i.e. subject vs. pattern based searches and,
2. Whether a new centralized and searchable database containing detailed query results would be created and available at the system hub?

Unfortunately and as detailed below, the full Committee resolved only the first conflict and declined to resolve the second.

The final report contains contradictory advice regarding the fundamentally critical issue of a new database.

Thus I voted "No" on the motion to send the combined report to the DHS staff.

1. Pattern vs. Subject Based Searches

Our Subcommittee was quite clear that our recommendations were contingent on the assumption that all searches would be subject based. We said that we would need to revisit our report if the ultimate system permitted pattern based searches.

In contrast, the Technical Subcommittee's report allowed for pattern based searches.

The Committee resolved this conflict in favor of our Policy Subcommittee by inserting language into the Technical section reiterating that the report was contingent on the assumption that all queries would be subject based using specific PII.¹

However, the Chief Privacy Officer stated in a public session of DPIAC that DHS had not ruled out using the federated system to conduct broad pattern based searches.

It is essential that we follow up with her on DHS's ultimate decision and revisit the report should DHS decide to conduct pattern based searches.

2. A New Database at the Hub

The second fundamental conflict -- what data would be kept at the hub after the queries were run -- was not resolved.

Consequently, the first section of the final report is drawn from the Policy Committee's recommendations:

“ One approach envisions a centralized database at the “hub,” which would contain pointers to participating component databases, the queries of users searching for information in other component databases, and the results of those queries. The hub also would contain an audit log. An alternative approach would be to retain far less information at the hub, limited to pointers to participating component databases and an audit log. The audit log would retain the queries and information on the users. “ (Final report page 3)

¹ “W al

This section incorporates the views of the Policy Subcommittee. We argued against the creation a new centralized database at the hub. We feared it would create a new set of privacy problems.

It is an essential component of our report which recognizes that DHS is a vast and complex agency. Its widely diverse components collect an extraordinary amount of PII. Much of that data is of extremely low quality, e.g. “suspicious activity reports” from the general public.

If query results were assembled into a database at the system hub, there would be no way to judge how accurate and reliable the underlying data was.

Many innocent Americans would be labeled as “suspect” based on their inclusion in the newly created database.

The second section of the report is drawn from the recommendation of the Technical Subcommittee.

In contrast to the first section incorporating the Policy Committee’s paper, this section is based on the polar opposite assumption that a full set of data – in effect a new centralized database that included query results—will be kept at the hub where it will be open to search and misuse.

The Technical Subcommittee proposed the language reproduced below, which was ultimately incorporated into the final report.

The Technical subgroup’s language presupposes that the search results will be kept at the hub.

The only qualifications are some rather generous and non-specific limitations on the length of time that data will be retained.

Privacy Technology Guidance

H. Data Retention

The data retention policies for the information sharing system should be predicated on the following two principles. 1) The actual queries (not the data retrieved there from) should be saved for the longest regulatory period, so that audit logs can be effective in understanding what people are querying and why; 2) The data inferred from those queries should be saved for the shortest regulatory period possible (essentially consistent with the reason why that query/data was assembled in the first place).

In seeming haste to conclude our discussion and without any substantive discussion, the Committee rejected a motion to reconcile the two approaches by including language clearly stating that we were recommending against the creation of a fulsome database at the hub.

The Technical Subcommittee's language was allowed to stand.

It appears in the final report at page 19.

The creation of a new database containing unreliable and unverified information about innocent Americans has profound implications for both privacy and data integrity.

I could not vote for a report with two sections that took such a fundamentally contradictory approach to this issue.

Thus my "No" vote on the motion to send the consolidated reports to the Department.

B. The Limitations of the Report

DPIAC has an extremely limited mandate. We are restricted to considering only those issues referred to us by the Department.

In this case, our charge from the Chief Privacy Officer² shaped both the content of the report and our discussions.

I cannot and will not discuss any matters that are not on the public record of the Committee.

But I will underscore what is plain from the publically available version of the document itself.

DPIAC's advice:

1. Assumed that the system would be built and does not address whether its construction was wise or necessary;
2. Does specifically discuss the predicate for a search/query. It does not set the bar for a search e.g. "reasonable suspicion" that the subject(s) has engaged in a crime or represents a threat;
3. Does not specifically address DHS' use of commercial data either as part of the proposed system or when it is combined with the query results and,
4. Does not clearly address the onward transfer of query results to outside parties such as Fusion Centers.

² Letter from Mary Ellen Callahan to Richard Purcell dated December 30, 2010