World Privacy Forum comments to the National Institute of Standards and Technology (NIST), US Department of Commerce, regarding the draft report on De-Identification of Personally Identifiable Information (NISTIR 8053)

Via email to draft-nistir-deidentify@nist.gov

Attn: Computer Security Division,
Information Technology Laboratory
100 Bureau Drive (Mail Stop 8930)
Gaithersburg, MD 20899-8930

May 13, 2015

The World Privacy Forum is pleased to have this opportunity to submit comments on the National Institute of Standards and Technology's draft report titled *De-Identification of Personally Identifiable Information* (NISTIR 8053), available at http://csrc.nist.gov/publications/drafts/nistir-8053/nistir_8053_draft.pdf.

The World Privacy Forum is a non-profit public interest research and consumer education group. We have published many research papers and policy comments focused on privacy and security issues. Much of our work explores technology and health-related privacy issues, biometrics, consent, data analytics, and many other rapidly evolving areas of privacy. You can see our publications and more information at www.worldprivacyforum.org.

The WPF welcomes the draft NIST report. The entire area of de-identification and re-identification of personal data swirls with controversy and confusion today. We see considerable value in a NIST report that describes terminology in this area, as well as considers the process and procedures for the removal of personally identifiable information (PII) from a variety of electronic document types.

Following are our specific comments on the draft report, identified by line item. Our substantive comments focus on IRBs and the HIPAA expert method of determining risk of re-identification.

I. Specific comments

A. Line 255: There appears to be a typo here, with the words *would result* likely superfluous.

B. Line 282: The words *same practitioners* are unclear. The word *same* could be read to refer back to the *different practitioners* on line 281, and that is not what we think is the intended reference. The sentences need revision and clarification.

C. Lines 381-385: The mention of the use of IRBs as potential regulators of the *other harms* is reasonable. However, while it is true that IRBs have a wide range of capabilities, some IRBs will not have the skills needed to identify and prevent the *other harms*, and IRBs approve a significant number of research protocols using expedited review procedures, with protocols that rely on de-identified information perhaps accepted for expedited review more often than those that use identifiable information. The mention of IRBs here without any comment or qualification seems to place more reliance on IRBs than is appropriate here. A sentence might be added to the effect that the capabilities of IRBs and other internal controls may not always be sufficient to provide the necessary review and oversight of the adequacy and consequences of the sharing of de-identified data.

D. Note 23: This note cites an article by Robert Gellman, but the citation is incomplete. A full citation is:

> *The Deidentification Dilemma: A Legislative and Contractual Proposal*, 21 Fordham Intellectual Property, Media & Entertainment Law Journal 33 (2010), http://iplj.net/blog/wp-content/uploads/2013/09/Deidentificatin-Dilemma.pdf.

E. Line 554: This key section opens up a discussion of HIPAA's expert method of determining that the risk of re-identification for any disclosure of de-identified data is minimal. The expert method has several problems that the NIST text does not mention. These are important problems that need to be surfaced in the report.

The first problem is the lack of transparency. The rule does not require that there be any public disclosure of the expert used, of the expert's qualifications, or of the methodology used by any given expert in any given case. While disclosure of methodology runs some risk of revealing business or research information that might quality for confidential treatment, disclosure of the hiring of an expert, the name of the expert and the expert's qualifications present no similar risk.

The second problem is that the standards for a qualified expert in the rule are minimal and vague. Any HIPAA-covered entity can accept the opinion of any "expert" that it chooses to recognize. Together with the lack of transparency, this opens the door to "expert shopping" or to the use of experts who have a conflict of interest or a known bias. Further, if a covered entity needs to hire a second expert to get the opinion that it desires, there is no way for the public to know that a previous expert refused to clear the disclosure.

The third problem is the likelihood of inconsistent results. The NIST report already reflects some of the disagreement in the expert community about the likelihood and significance of re-identification of de-identified data. Given the wide range of current opinion among experts, any covered entity or data user can readily find the opinion it wants. There is a significant risk of a race to the bottom here.

The fourth problem is the lack of any meaningful oversight by the Office of Civil Rights. That office is already overwhelmed with general oversight of the rule and with reports of data breaches. OCR is not likely to have the luxury of reviewing expert determinations. The rule does not even require reporting of the fact that a determination was made or even that a negative

determination preceded a positive determination. In addition, anyone adversely affected by re-identification of information disclosed on the basis of an expert's opinion is unlikely to know of the cause and may have some difficulty knowing where to report the concern.

We do not expect the NIST report to offer a complete critique of the HIPAA expert methodology. We agree that there may be an appropriate role for expert determinations under HIPAA and otherwise. However, we think that the report should include some comment about the shortcomings of the expert method, including the lack of transparency, lack of standards for qualified experts, possibility of inconsistent results from the use of an expert methodology, and shortcomings of oversight.

II. Conclusion

Again, we appreciate the draft report on de-identification. It is important and timely. But we would like to see the known problems with both the IRB and HIPAA expert method of protections discussed with more clarity so as to carefully assess existing shortcomings and improve solutions in this area. We believe improvements can and should be made, rendering these systems more effective.

We thank you for the opportunity to submit these comments.

Respectfully,

Pam Dixon
Executive Director,
World Privacy Forum