VIA THE WEB

Federal Trade Commission
600 Penn. Ave. NW
Washington, DC 20580

Re: Request to Participate in PrivacyCon, January 14, 2016

Dear PrivacyCon Organizers

I would be delighted to present research on deidentification at the January 14, 2016 PrivacyCon event. I recently completed this joint research with Professor Woodrow Hartzog of Samford University's Cumberland School of Law and we are now in the process of publishing the attached paper, entitled *Anonymization and Risk*, in the Washington Law Review.

Our work confronts the failure of privacy law to grapple with the acknowledged limitations of anonymization. We argue that the debate over deidentification has stagnated and that the best way to advance debate is by focusing on the process of minimizing the risks associated with releasing data rather than on "anonymized" outputs.

The law has been slow to adopt a holistic approach to protecting data subjects when data sets are released to others. Currently, the law is focused on whether an individual can be identified within a given set. But data scientists regularly demonstrate that the concept of perfect anonymity is a pipe dream. The debate over how to protect people within data sets has been intense, yet has failed to result in much meaningful change. The law is like an ostrich with its head in the ground.

There is a better path forward. We argue that data release policy should pivot to a focus on the process of minimizing the risk of reidentification and sensitive attribute disclosure. By focusing on process, data release policy can mimic data security policy and better balance privacy and utility where nearly all data exchanges carry some risk.

In terms of the three points you asked applicants to address:

1. As to the *findings* we plan to present: We build upon the FTC's own process-based approach to protecting data subjects by showing how the full spectrum of techniques from the field of statistical disclosure limitations (SDL) can be used to tailor data release obligations to risk. In broad terms, we argue that the law of data release should look more like the law of data security: process-based, contextual, and tolerant of harm, so long as procedures to minimize risk are implemented *ex ante*. More specifically, we identify and discuss seven risk vectors that must be considered prior to any data release: data volume, data sensitivity, type of data recipient, data use, data treatment technique, data access controls, and consent and consumer expectations. Finally, we propose several legal reforms to implement process-based data release policy, including a general requirement for "reasonable" data release protections and a prohibition on deceptive deidentification.

2. As to our *methodology*, we reviewed in great depth not only the technical literature on deidentification but also the three major forms of SDL: direct access, dissemination-based access (including deidentification), and query-based access (including differential privacy). In analyzing these different methods, we assumed that none of them are perfect in all settings but that each method has something of value to contribute. In other words, we avoided taking sides in what we described as a debate between "formalists" (for whom mathematical proof is the touchstone of any meaningful policy) and "pragmatists" (for whom workable solutions should prevail over theoretical concerns). We instead sought a more policy-driven, integrated, and comprehensive approach to data release that would better protect data while preserving its utility.
3. As to *how our research differs from prior research in the area*: First, we looked at the SDL literature in a comprehensive manner and found that the two main camps were deeply divided in their goals, methods, interests, and measures of success, and rarely came together to share ideas and insights. Second, our paper is among the very first to argue that from a policy standpoint, exaggerating the merits or demerits of one side or the other is misguided, and that it is far more productive to figure out where and how the two sides come together. Finally, we stake out the position, which is rare in the literature, that the full spectrum of possible data release protections should be utilized to tailor a company's obligations to the likely level of risk, and we set forth an approach that may be widely adopted by regulators across multiple sectors.

Please note that an earlier version of our paper has been recognized within the international privacy community. Participants of the 8th Annual Privacy Law Scholars Conference selected our draft article for an encore workshop session. Our paper was also competitively selected for presentation at the upcoming 2015 Amsterdam Privacy Conference.

The attachment included both an abstract and the full text of our paper. I hope that you will grant our request to present our research at PrivacyCon.

Sincerely,

Ira Rubinstein