Dr. Urs Gasser
Executive Director, Berkman Center for Internet & Society, Harvard University
Professor of Practice, Harvard Law School
23 Everett St, 2nd Floor
Cambridge, MA 02138

October 9, 2015

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

Re: Request for Research Presentations for the PrivacyCon Conference

Dear Kristen Anderson and Daniel Salsburg,

I respectfully submit this proposal to deliver a presentation at the PrivacyCon conference to be held by the Federal Trade Commission on January 14, 2016. Through Harvard University's Privacy Tools for Sharing Research Data collaboration, my colleagues and I have been studying the issues that arise when collecting, analyzing, and disseminating data containing personal information. These efforts are focused on translating the theoretical promise of new measures for privacy protection and data utility into practical tools, approaches, and recommendations.

A portion of this research is summarized and applied in the attached article, "Towards a Modern Approach to Privacy-Aware Government Data Releases," which will be published in the *Berkeley Technology Law Journal* this fall. In this article, we analyze data sharing models and methods for protecting privacy that are in common use. We find that traditional approaches to privacy are largely ad hoc and mismatched to the privacy risks and intended uses of data. In addition, the practices used are quite narrow compared to the wide range of privacy interventions available. In response, we propose a more systematic framework for privacy analysis and discuss how it can be used to select appropriate privacy controls for specific data collection, storage, use, and sharing cases.

While the article draws primarily from examples found in releases of government data, the findings and proposed framework are more broadly applicable. In particular, we believe this approach is also relevant to analyzing and guiding commercial privacy practices. Members of our research team will be presenting this work and addressing its applicability to the consumer privacy context at the upcoming *Conference on Responsible Use of Open Data: Government and the Private Sector* organized by the Berkeley Center for Law & Technology and NYU's Information Law Institute and Department of Media, Culture and Communication.

Please find below the relevant contact information, a full abstract for the article, and descriptions of key findings, the methodology, and a statement on how this research differs from prior research in this area.

Contact information

Dr. Urs Gasser
Executive Director, Berkman Center for Internet & Society, Harvard University
Professor of Practice, Harvard Law School
23 Everett St, 2nd Floor
Cambridge, MA 02138
ugasser@cyber.law.harvard.edu
(617) 495-7547

Article abstract

Towards a Modern Approach to Privacy-Aware Government Data Releases, by Micah Altman, Alexandra Wood, David R. O'Brien, Salil Vadhan, and Urs Gasser

Governments face increasing pressure to promote transparency, accountability, and innovation by making the data they hold available to the public. Because much of the data pertains to individuals, agencies rely on various standards and interventions to protect privacy interests while supporting a range of beneficial uses of the data. However, there are growing concerns among privacy scholars, policymakers, and the public that these approaches are incomplete, inconsistent, and difficult to navigate.

To uncover gaps in current practice, this article examines releases of data in responses to freedom of information and Privacy Act requests, traditional public and vital records, official statistics, and e-government and open government initiatives. It finds that agencies lack formal guidance for choosing among and implementing privacy interventions in specific cases. Most agencies address privacy by withholding or redacting records that contain directly or indirectly identifying information based on an ad hoc balancing of interests, and similar privacy risks are sometimes treated vastly differently by different government actors. These observations demonstrate the need for a more systematic approach to privacy analysis and suggest an outline for a new way forward.

In response, this article proposes a framework for a modern privacy analysis informed by recent advances in data privacy from disciplines such as computer science, statistics, and law. Modeled on an information security approach, this framework distinguishes between and characterizes privacy controls, threats, vulnerabilities, and utility. When developing a data release mechanism, policymakers should specify the desired data uses and expected benefits, examine each stage of the data lifecycle to identify privacy threats and vulnerabilities, and select controls for each lifecycle stage that are consistent with the uses, threats, and vulnerabilities at that stage. The article sketches the contours of this analytical framework, populates selected portions of its contents, and illustrates how it can inform the selection of privacy controls by discussing its application to two real-world examples of government data releases.

Findings

- Governments use a narrow set of tools to analyze and mitigate privacy risks, despite the broad range of privacy interventions proposed by privacy scholars, legal scholars, non-profit organizations, and many others. Most agencies address privacy concerns in the same fashion: by withholding or redacting records that contain certain pieces of directly or indirectly identifying information.
- Data release decisions are often based on an ad-hoc balancing of interests or the redaction of certain fields deemed to be directly or indirectly identifying information, despite evidence from the privacy science literature that such approaches will likely fail to address privacy risks. In addition, these decisions often result in the withholding useful information that could be safely shared.
- Guidance on interpreting and applying regulatory standards for privacy protection is remarkably thin. General guidance on protecting the privacy of individuals and preventing the release of personally identifiable information is available, yet there is relatively little regulatory guidance for formally characterizing privacy risks and selecting and implementing interventions in specific settings.
- The treatment of data across actors is largely inconsistent. Similar privacy risks—and, in some cases, even identical sets of data—are addressed differently by different government actors.
- In the rapidly changing environment of information policy and technology, neither science nor principle provides definitive guidance on how to select policy components for a data release based on the risks and benefits of each case.
- An information lifecycle framework, while not yet fully prescriptive, can provide a systematic and useful decomposition of the factors relevant to data release, and can be used to order the set of interventions that should be considered at each stage.
- Changes in science and technology offer the opportunity for sophisticated characterization of privacy risks and harms, as well as modern forms of educational interventions and technical controls. Policymakers now have the opportunity to select from a distinct set of legal, technical, economic, procedural and educational interventions at each stage, in order to construct a comprehensive policy that is based on the specific uses, threats, and vulnerabilities of the release.
- A systematic framework such as the one we propose provides a natural foundation for increased transparency, through the documentation of the uses, potential risks, and the privacy and security interventions selected at each lifecycle stage.

Methodology

As described in detail in the accompanying draft article, our findings are based on a literature review, use case analysis, and series of expert interviews. Our analysis recognizes insights from the computer science, social science, statistics, law and policy, and information science literature. In particular, we draw from research on quantitatively measuring privacy and utility, assessing the effectiveness of legal and policy governance mechanisms, and balancing research opportunity and the nature of privacy risks. We also employ information security approaches to cataloging and characterizing data uses, privacy interventions, and threats and vulnerabilities related to privacy,

and an information lifecycle model for tracing actors and actions on information from collection and retention to release and post-access.

In addition, to gather information about current data sharing models and privacy practices, we engaged in a use case analysis of four different categories of real-world data releases. For each use case, we analyzed the actors involved; the types of information released; the legal requirements and institutional policies for making release decisions; and the legal, technical, and procedural approaches to privacy in use. We supplemented this review with expert interviews with city open data managers who assess data privacy risks, make data release decisions, and implement legal, technical, and procedural privacy interventions.

How this research differs from prior research in the area

Our analysis differs from prior research primarily with respect to the depth of its engagement with the technical literature on privacy, its interdisciplinary methodology, and the resulting framework for privacy analysis. The draft article provides a catalog of the range of procedural, economic, educational, legal, and technical interventions for protecting privacy, as well as commentary explaining advances in these areas to a general audience. Drawing from information security approaches and a lifecycle analysis model, it proposes a new framework for privacy analysis. Finally, it illustrates how organizations that manage data can use such a framework to analyze privacy risks more systematically and apply concepts from recent advances from the multidisciplinary privacy literature in real-world settings.

Thank you very much for your consideration.

Sincerely,

Urs Gasser