epic.org

Electronic Privacy Information Center
1718 Connecticut Avenue NW, Suite 200
Washington, DC 20009, USA

📞 +1 202 483 1140
🖨 +1 202 483 1248
🐦 @EPICPrivacy
🌐 https://epic.org

July 11, 2018

The Honorable Randy Weber, Chair
The Honorable Marc Veasey, Ranking Member
The Honorable Barbara Comstock, Chair
The Honorable Daniel Lipinski, Ranking Member
Subcommittee on Energy & Subcommittee on Research and Technology
House Committee on Science, Space, and Technology
2321 Rayburn House Office Building
Washington, DC 20515

Dear Chair Weber, Chair Comstock, Ranking Member Veasey, and Ranking Member Lipinski:

We write to you regarding the upcoming hearing on "Big Data Challenges and Advanced Computing Solutions."[1] The Electronic Privacy Information Center ("EPIC") strongly supports efforts to make data in the federal government more widely available to ensure better policymaking. At the same time, where data maintained by the federal government implicates identifiable individuals, privacy risks must be addressed and reduced as much as possible. Privacy should not be sacrificed for the sake of evidence-based policymaking, and, if certain safeguards are put in place, evidence-based policymaking should not need to be sacrificed for the sake of privacy.[2]

EPIC is a public-interest research center established in 1994 to focus public attention on emerging privacy and civil liberties issues. EPIC has long advocated for privacy and security safeguards for data as well as the use of privacy enhancing technologies ("PETs") that minimize or eliminate the collection of personally identifiable information.[3] EPIC testified before the Commission on Evidence-Based Policymaking and called for the Commission to adopt innovative privacy safeguards to protect personal data and make informed public policy decisions.[4] Additionally, EPIC President Marc Rotenberg and EPIC Advisory Board member Cynthia Dwork served on a panel at the National Academies of Science that released a report on how federal data sources can be used for public policy research while protecting privacy.[5]

---

[1] *Big Data Challenges and Advanced Computing Solutions*, 115th Cong. (2018), H. Comm. on Science, Space, and Technology, https://science.house.gov/legislation/hearings/subcommittee-energy-and-subcommittee-research-and-technology-hearing-big-data.

[2] EPIC Executive Director Marc Rotenberg, Let's Use Government Data to Make Better Policy, Scientific American (Oct. 4, 2017), https://blogs.scientificamerican.com/observations/let-rsquo-s-use-government-data-to-make-better-policy/.

[3] *See, e.g.*, EPIC Executive Director Marc Rotenberg, Testimony Before the U.S. House of Representatives Committee on Energy and Commerce, Subcommittee on Commerce, Trade, and Consumer Protection, Mar. 1, 2001, *Privacy in the Commercial World*, https://epic.org/privacy/testimony_0301.html.

[4] Marc Rotenberg, Commission on Evidence-Based Policymaking: Privacy Perspectives, before the National Academies of Science, Sep. 9, 2016, https://epic.org/privacy/wiretap/RotenbergCEBP-9-16.pdf.

[5] National Academies of Science, "Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy" (2017), https://www.nap.edu/catalog/24652/innovations-in-federalstatistics-combining-data-sources-while-protecting-privacy.

Requiring comprehensive risk assessments for de-identified confidential data and supporting adoption of PETs are key to protecting personal information. Even where data has been de-identified it is still possible to combine certain data sets with others to determine extensive amounts of personal information.[6] Moving forward, government agencies conducting evidence-based policymaking should adopt PETs to reduce the risk of re-identification of personal data. As was noted in the report by the National Academies of Science:

> Any consideration of expanding data must have privacy as a core value…As federal agencies seek to combine multiple datasets, they need to simultaneously address how to control risks from privacy breaches. Privacy-enhancing techniques and privacy-preserving statistical data analysis can be valuable in these efforts and enable the use of private-sector and other alternative data sources for federal statistics.[7]

Equally important is to recognize that under the Privacy Act statistical data is subject to fewer privacy constraints because it is understood that statistical does not identify specific individuals. If it is possible to re-identify aggregate data, complete privacy protections must necessarily apply. Agencies will carry the responsibility to ensure the adequacy of the privacy enhancing and privacy protecting techniques.

We ask that this statement be entered in the hearing record. EPIC looks forward to working with the Committee on these issues of vital importance to the American public.

Sincerely,

/s/ *Marc Rotenberg*
Marc Rotenberg
EPIC President

/s/ *Christine Bannan*
Christine Bannan
EPIC Administrative Law and Policy Fellow

---

[6] Latanya Sweeney, Simple Demographics Often Identify People Uniquely, Carnegie Melon University, Data Privacy Working Paper, 2000, https://dataprivacylab.org/projects/identifiability/paper1.pdf
[7] Innovations in Federal Statistics at 3.