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

May 2, 2017

The Honorable John Culberson, Chairman
The Honorable Jose Serrano, Ranking Member
U.S. House of Representatives Committee on Appropriations
Subcommittee on Commerce, Justice, Science, and Related Agencies
H-305, The Capitol
Washington, DC 20515

Dear Chairman Culberson and Ranking Member Serrano:

We write to you regarding the hearing on "Oversight of the 2020 Census."[1] EPIC is a public interest research center established in 1994 to focus public attention on emerging privacy and civil liberties issues. The Census implicates numerous privacy issues.[2] As the Subcommittee evaluates the 2020 Census, there are several issues that we ask you to consider.

The Census is an essential part of understanding the changing demographics in America. The census helps ensure evidence-based policy decisions and census data is the source of much political and economic planning in the United States. However, it is of the utmost importance the individual privacy is respected. Every effort must be taken to ensure that the personal information of individuals and that census data is not used improperly.

As technology has advanced, it has become easier to obtain the identity of an individual from aggregate information. Dr. Latanya Sweeney demonstrated this in experiments she conducted using the results of the 1990 Census. Dr. Sweeney determined that 87% of the population had reported characteristics that made it possible to personally identify them.[3] Furthermore, she determined that half of the population could be identified based on the city or town they lived in, gender, and date of birth.[4]

Dr. Sweeney was again able to show that anonymous data may not in fact be anonymous when she was able to identify former Massachusetts Governor William Weld based on the release of aggregate data of Massachusetts state employees.[5]

---

[1] *Oversight of the 2020 Census*, 115th Cong. (2017), S. Comm. on Commerce, Justice, Science, and Related Agencies, http://appropriations.house.gov/calendar/eventsingle.aspx?EventID=394839 (May 3, 2017).

[2] *The Census and Privacy*, EPIC, https://epic.org/privacy/census/.

[3] Latanya Sweeney, *Simple Demographics Often Identify People Uniquely*, Carnegie Melon University, Data Privacy Working Paper 3, 2000, https://dataprivacylab.org/projects/identifiability/paper1.pdf.

[4] *Id.*

[5] Nate Anderson, *"Anonymized" Data Really Isn't---And Here's Why Not*, Arts Technica, Sept. 8, 2009, https://arstechnica.com/tech-policy/2009/09/your-secrets-live-online-in-databases-of-ruin/.

There is also reason to be concerned about how Census data could be used for improper purposes. During World War II, the Census Bureau provided information to the War Department that led to the internment of 120,000 innocent Japanese.[6]

Given the concerns about the potential misuse of census data after 9-11, EPIC pursued a Freedom of Information Act request. Documents from the 2000 Census obtained by EPIC revealed that the Census Bureau had provided the Department of Homeland Security with data on individuals of Arab ancestry.[7] As we explained in 2004:

> EPIC obtained documents through the Freedom of Information Act last week revealing that the Census Bureau gave the Department of Homeland Security statistical information on people who identified themselves on the 2000 census as being of Arab ancestry. The special tabulations were prepared specifically for the law enforcement agency, and do not indicate that similar information about any other ethnic groups was requested. The tabulations apparently include information about United States citizens, as well as individuals of Arab descent whose families have lived in the United States for generations.
>
> One tabulation shows cities with populations of 10,000 or more and with 1,000 or more people who indicated they are of Arab ancestry. For each city, the tabulation provides total population, population of Arab ancestry, and percent of the total population which is of Arab ancestry.
>
> A second tabulation, more than a thousand pages in length, shows the number of census responses indicating Arab ancestry in zip codes throughout the country. The responses indicating Arab ancestry are subdivided into Egyptian, Iraqi, Jordanian, Lebanese, Moroccan, Palestinian, Syrian, Arab/Arabic, and Other Arab.
>
> The heavily redacted documents show that in April 2004, a Census Bureau analyst e-mailed a Department of Homeland Security official and said, "You got a file of Arab ancestry information by ZIP Code Tabulation Area from me last December (2003). My superiors are now asking questions about the usage of that data, given the sensitivity of different data requests we have received about the Arab population.[8]

---

[6] W. Seltzer and M. Anderson, "Census Confidentiality under the Second War Powers Act (1942-1947)." Paper prepared for presentation at the annual meeting of the Population Association of America, New York, March 29-31, 2007, *available at* http://studylib.net/doc/7742798/census-confidentiality-under-the-second-war-powers.

[7] *Department of Homeland Security Obtained Data on Arab Americans From Census Bureau*, EPIC, https://epic.org/privacy/census/foia/; Lynette Clemetson, *Homeland Security Given Data on Arab-Americans*, New York Times, Jul. 30, 2004, http://www.nytimes.com/2004/07/30/us/homeland-security-given-data-on-arab-americans.html.

[8] *EPIC Report: Homeland Security Given Census Data on Arab Americans,* EPIC Alert 11.15 (Aug. 4, 2005), https://epic.org/alert/EPIC_Alert_11.15.html

EPIC's FOIA efforts led the Census Bureau to implement new procedures regarding the release of "potentially sensitive data to requesting law enforcement agencies and organizations or individuals."[9] EPIC supported those efforts but recognizes also the ongoing concerns about the potential misuse of data collected by government agencies. Amid rising fears that some minority groups may be targeted by law enforcement agencies, your committee should ensure that the data collected by the federal government is not misused.[10]

We note also that the National Academies of Science is currently pursuing a study "Using Multiple Data Sources and State-of-the-Art Estimation Methods in Federal Statistics: Frameworks, Methods, and Assessment" that may be of interest to the committee.[11]

EPIC looks forward to working with the Subcommittee to ensure that the census data provides the maximum benefit to the American public while minimizing the privacy risks.

We ask that this letter from EPIC be entered in the hearing record.

Sincerely,

/s/ Marc Rotenberg

Marc Rotenberg
EPIC President

/s/ Caitriona Fitzgerald

Caitriona Fitzgerald
EPIC Policy Director

/s/ Kim Miller

Kim Miller
EPIC Policy Fellow

---

[9] Census Bureau News, ''U.S. Census Bureau Announces Policy Regarding Sensitive Data,'' press release CB04-145, August 30, 2004; Lynette Clemetson, *Census Policy On Providing Sensitive Data Is Revised*, New York Times, Aug. 31, 2004, http://www.nytimes.com/2004/08/31/us/census-policy-on-providing-sensitive-data-is-revised.html.
[10] Tal Kopan, *Trump Admin Sets Stage For Mass Deportations,* CNN, Feb. 22, 2017, http://www.cnn.com/2017/02/21/politics/dhs-immigration-guidance-detentions/.
[11] *See, e.g.,* National Academies of Science, "Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy (2017), https://www.nap.edu/catalog/24652/innovations-in-federal-statistics-combining-data-sources-while-protecting-privacy