ENSURING CONFIDENTIALITY IS AT THE HEART OF DIFFERENTIAL PRIVACY, A NEW DATA PROTECTION SYSTEM BEING IMPLEMENTED FOR THE 2020 CENSUS.

AND SINCE THIS WILL BE THE FIRST DECENNIAL CENSUS PRIMARILY TAKING PLACE ONLINE, THERE’S ADDED CONCERN ABOUT THE SECURITY OF THE DATA BEING COLLECTED.

CENSUS RESULTS FORM A VAST DATABASE OF INFORMATION ABOUT US AS A COUNTRY...

THE ANONYMIZED PATTERNS OF DEMOGRAPHIC GROUPS COLLECTED VIA THE CENSUS IS AN INVALUABLE PUBLIC RESOURCE.

...BREAKDOWNS BY GENDER, AGE, AND RACE/ETHNICITY; HOUSEHOLD SIZE; RENTERS VERSUS HOME-OWNERS; AND REGIONAL POPULATION DENSITY.
IF SOMEONE COULD HACK THE CENSUS, IT COULD BE CRITICALLY DAMAGING TO OUR COUNTRY, ESPECIALLY IF THEY RELEASED THAT DATA AS PART OF A DISINFORMATION CAMPAIGN.

SITUATIONS LIKE THE CAMBRIDGE ANALYTICA DATA SCANDAL HAVE RAISED CONCERNS THAT A BAD ACTOR COULD THEORETICALLY TAKE THE AGGREGATE ANONYMOUS INFORMATION...

...AND REVERSE-ENGINEER CENSUS DATA BY COMBINING IT WITH OTHER AVAILABLE RECORDS...

...TO UNMASK THE PERSONAL IDENTITIES AND/OR PRIVATE INFORMATION OF ANYONE WHO PARTICIPATED IN THE CENSUS.

IF SOMEONE COULD HACK THE CENSUS, IT COULD BE CRITICALLY DAMAGING TO OUR COUNTRY, ESPECIALLY IF THEY RELEASED THAT DATA AS PART OF A DISINFORMATION CAMPAIGN.

THE CAPACITY FOR VOTER SUPPRESSION WOULD CERTAINLY EXIST WITH A MAJOR BREACH.

THE CENSUS BUREAU COLLECTS INFORMATION ABOUT RACE, GENDER... SOMEONE COULD EASILY SEGMENT OR SLICE UP THE ADULT POPULATION AND USE THAT FOR COMMERCIAL OR POLITICAL PURPOSES.

THAT’S WHY THE CENSUS BUREAU IS IMPLEMENTING DIFFERENTIAL PRIVACY FOR THE 2020 CENSUS.

DIPAYAN GHOSH, PRIVACY EXPERT AND SENIOR RESEARCH FELLOW AT HARVARD KENNEDY SCHOOL’S SHORENSTEIN CENTER
How does Differential Privacy work? Geographically, small-area Census data is reported in "Census Blocks"—generally equivalent in size to a city block.

Each Census block provides basic characteristics about the residents: age ranges, gender breakdowns, and race/ethnicity information.

Census microdata—which includes household size and the relationships between residents—is derived from the individual information collected by the Census Bureau, and is reported in groups of at least 100,000 people.

With Differential Privacy the Census Bureau aims to release all that aggregate information—and still protect individual privacy.

By introducing "noise" into the data, Differential Privacy masks the private information of the individuals within the database.

The challenge with implementing Differential Privacy is maintaining a balance...

It's a technical trade-off between two things--

-- understanding more about the information in a database and maintaining the privacy of those users.

Accuracy
Privacy loss
Confidentiality
Accuracy loss

(Noise is a controlled quantity of random data.)

The 2010 U.S. Census included 11 million blocks (of which 6.3 million were inhabited).