DIFFERENTIAL PRIVACY

A CONFIDENTIALITY PROTECTION PLAN FOR THE 2020 CENSUS

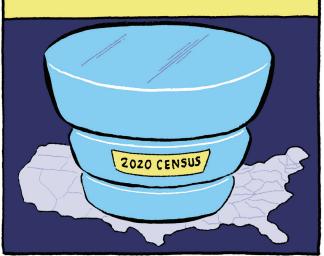
By JOSH NEUFELD



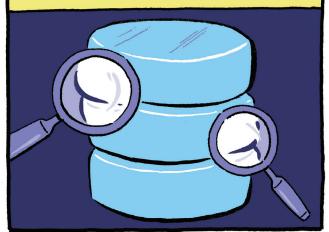
AND SINCE THIS WILL BE THE FIRST DECENNIAL CENSUS PRIMARILY TAKING PLACE **ONLINE**, THERE'S ADDED CONCERN ABOUT THE **SECURITY** OF THE DATA BEING COLLECTED.



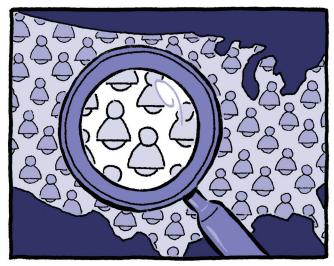
CENSUS RESULTS FORM A VAST DATABASE OF INFORMATION ABOUT US AS A COUNTRY...

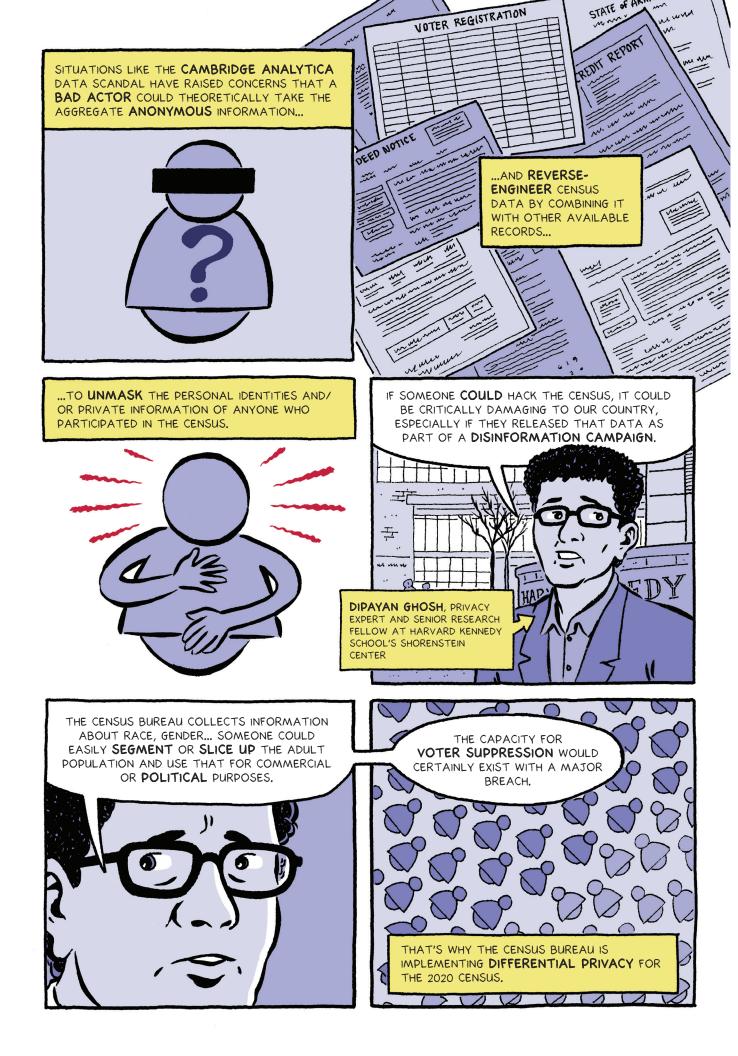


...BREAKDOWNS BY GENDER, AGE, AND RACE/ ETHNICITY; HOUSEHOLD SIZE; RENTERS VERSUS HOME-OWNERS; AND REGIONAL POPULATION DENSITY.

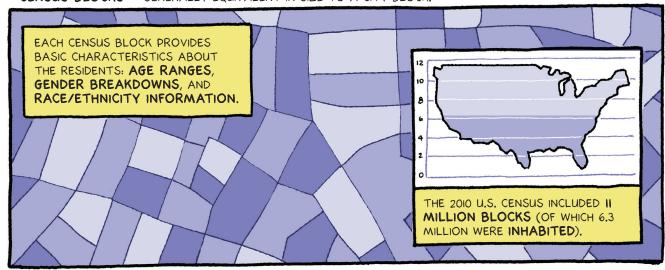


THE ANONYMIZED PATTERNS OF DEMOGRAPHIC GROUPS COLLECTED VIA THE CENSUS IS AN INVALUABLE PUBLIC RESOURCE.

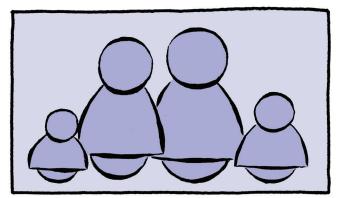




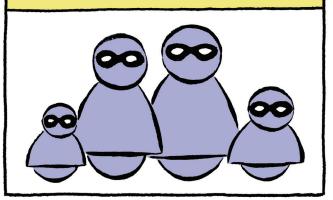
HOW DOES DIFFERENTIAL PRIVACY WORK? GEOGRAPHICALLY, SMALL-AREA CENSUS DATA IS REPORTED IN "CENSUS BLOCKS" -- GENERALLY EQUIVALENT IN SIZE TO A CITY BLOCK.



CENSUS MICRODATA -- WHICH INCLUDES HOUSE-HOLD 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.

ACCURACY
PRIVACY LOSS

(NOISE IS A CONTROLLED QUANTITY OF RANDOM DATA.)

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.

