

Stanford Report, February 2, 2010

Count on this: Stanford researchers get better access to census data

The Secure Data Center, protected by several privacy protocols, will let researchers tap detailed census information that could serve as the framework for public policy.

BY ADAM GORLICK

L.A. Cicero



Sociology Professor C. Matthew Snipp; Chris Thomsen, executive director of IRISS; and Professor Nick Bloom, associate professor of Economics

If you're keeping count of the research facilities at Stanford University, the population just increased by one. And this new kid on the Farm is related to the granddaddy of number crunchers, data collectors and all things countable – [the U.S. Census Bureau](#).

Thanks to an arrangement with the university's Institute for Research in the Social Sciences (IRiSS), a Secure Data Center is opening on campus that will allow faculty and student researchers access to an ocean of confidential information accumulated by the Census Bureau, the National Center for Health Statistics and other federal agencies.

While the government releases troves of demographic and economic facts and figures to the public, academic researchers are allowed additional access, to tap the nitty-gritty details about individuals, communities and private companies that are otherwise restricted by the government.

Most demographic information publicly released by the Census Bureau is organized in tracts, which represent large geographic areas and as many as 25,000 people. In order to get a snapshot of a smaller area – like a few city blocks – researchers need access to the secure data. So someone studying how a neighborhood's racial makeup coincides with its educational performance, for instance, would have to aggregate information that can only be obtained from the secured data.

"The idea is that the government could have people at places like Stanford who are dedicated to sifting through and analyzing all this information," said C. Matthew Snipp, faculty director of the Secure Data Center and director of Stanford's Center for Comparative Studies on Race and Ethnicity. "That research gets used as a basis for policymaking. But it's hard to do that kind of analysis in the federal system because of all the other demands they have."

Not to worry: Privacy is protected by a series of security measures

Before getting access to the secure data, academics must have their research projects approved by the government – a review process that examines not only the intellectual merit of the research but also the research benefit to the federal government and the necessity of accessing the secure data. Once they get the OK and pass a background investigation, the scholars sit in a rather unremarkable office guarded by multiple alarms and under constant watch of security cameras.

A terminal accesses the Census Bureau's servers in Maryland; no data is stored locally. While researchers may be looking at sensitive data collected by the government – such as personal income, religious

affiliation and health statistics – they're not allowed to take any notes.

If they want to print anything out, they have to tabulate the information and scrub it of any personal or private data. Those documents are sent to a printer in an office next door, where an administrator from the Census Bureau reviews them to make sure there are no disclosure issues, then hands the printouts to the researcher.

Any breaches in security or protocol are subject to heavy fines and prison terms.

"They take security and privacy very, very seriously," said Chris Thomsen, executive director of IRISS. "Keeping the public's trust is vitally important."

The Secure Data Center at Stanford is a satellite operation of a California Census Research Data Center located at the University of California-Berkeley. The only other CCRDC is hosted by UCLA.

Social scientists have long been able to access secure data from the Census Bureau, but it wasn't always possible unless they had the luxury of time and money to go either to the agency's Maryland headquarters or one of the Research Data Centers that came online about 10 years ago.

"Having to go to Berkeley was an immense waste of time," said Nick Bloom, an associate professor of economics focused on what caused the recent recession and the reasons behind the recovery. "Being able to access this information here will be a great help. And it will be easier to have graduate students work with this data."

For Bloom, the census data holds important clues to economic performance that couldn't otherwise be charted and analyzed. Most American companies are privately held, and information about them isn't publicly released. So without the secure data, with its company details, Bloom and other social scientists would have an incomplete sense of the country's corporate landscape.

"If you want truly representative data, you need the census," he said. "Without it, you're only getting a sliver of the story."

MEDIA CONTACT

Adam Gorlick, Stanford News Service: (650) 725-0224,
agorlick@stanford.edu

© Stanford University. Stanford, California 94305.
[Copyright Complaints](#) [Trademark Notice](#)