For the 2020 Census, A Data-Driven Approach Counts

New technologies are helping the U.S. Census Bureau reduce costs and encourage self-response.

James Mersol

Thu, 03/05/2020 - 12:42

Photo Credit: wildpixel/iStock

The goal of the 2020 Decennial Census is to count every resident "once, only once and in the right place," as the U.S. Census Bureau's mission states. With an estimated 329 million residents living in the U.S. — though census itself will show the exact number – this is the largest undertaking in U.S. history.

The census is constitutionally mandated, of course, but also plays a critical role in informing how over $675 billion in government appropriations are allocated, how many representatives each state gets and how voting precinct and school district lines are drawn. Census data also helps local, state and federal agencies as well as companies make important decisions going forward.

"We need that measurement in order to shape our future," said Atri Kalluri, senior advocate for response security and data integrity at the U.S. Census Bureau, speaking at the Splunk Gov Summit March 4.

The census could not occur at this scale, nor within its mandated timeframe, without a data-driven approach to innovation.

One way the Census Bureau innovated was in how it reengineered address canvassing. Before every decennial census, the bureau validates addresses, looking for new housing units from the last census as well as removing any that no longer exist from its database. This time around, Kalluri explained, the bureau was able to verify 65% of housing units across the U.S. from its office, using a combination of imaging and change detection technology. For the remaining 35%, the bureau hired 32,000 assistants to verify addresses in the field, compared to the 150,000 it hired in 2010.

Data-enabled platforms and automation are also essential to the process of conducting the census, Kalluri said. This year, the Census Bureau is looking to hire up to 500,000 part-time employees to work as field workers. Not only does such a massive short-term increase in hiring, training and payroll necessitate a fast-moving IT platform, but also each one of these field operatives needs a device to conduct their operations. The Census Bureau has adopted a "device as a service" model to drive down equipment costs. These devices are used both for enumeration purposes and to optimize routes for enumerators, allowing them to survey more households in a shorter period of time.

The Census Bureau has implemented a system of checks and balances to ensure the accuracy of integrity of its data. Every response, whether with an ID or without (such as someone who fills out the census on a mobile device) is tested against rules and logic. For example, if someone reports that they are older than their parents, the system will bounce that response back. The system uses a variety of data sets to verify addresses as well; for the few that the system cannot automatically process, the bureau will use field verification.

The bureau has founded a trust and safety team, which is coordinating with multiple tech companies and non-governmental organizations to protect against misinformation and disinformation. Some of the social media organizations the trust and safety team has met with have changed their policies on posting and allowing bots to publish media, helping to stem the flow of false information.

The trust and safety team at the Census Bureau is mitigating misinformation and disinformation efforts across foreign languages as well, an especially important mission as the agency expands its foreign language offerings. This year, respondents to the census have the option to self-respond online in 13 different languages (including English). The paper self-response has guides in 59 different languages, and the Census Bureau's website will soon include text and video guides in those languages.

The Census Bureau's language choices were informed by the American Community Survey, a monthly survey the bureau sends out to a sample of the population, Kalluri said. Similarly, the agency used existing studies on response methods in its decision to add an online self-response option this year. It predicts 60.5% of respondents will self-respond.

The Census Bureau is committed to "very strictly maintain the confidentiality and privacy of data" as well as data integrity, Kalluri emphasized. Not only is unique respondent data not publicly available, but also the bureau has taken steps to protect the aggregate data from the sort of reverse-engineering that would allow someone to identify a specific respondent or group of respondents.

"Every employee and contractor at the Census Bureau is sworn for life to protect the data's integrity," Kalluri said.

The 2020 Decennial Census period officially starts on Census Day, April 1, Kalluri said, although some counts in difficult-to-reach communities in Alaska began in January. Some U.S. residents may also receive information as soon as March 12, and everyone is encouraged to self-respond as soon as they can.

[View printer friendly version](#)
[Census Bureau](#)
[Census](#)
[data integrity](#)
[data](#)
[Standard](#)