

Planned Innovations in the United States 2020 Census¹

Robert Colosi

United States Census Bureau, Washington, DC – United States of America, USA

E-mail: Robert.W.Colosi@Census.Gov

Abstract. In the United States of America, the Census Bureau (USCB) is implementing four major areas of innovation in preparation for the 2020 Census. The first area focuses on the maintenance of our frame of addresses using satellite imagery. The USCB will compare old and new satellite images to identify geographic areas that are stable and will not require in-person field staff updating. The second uses the Internet, including Identified (ID) and Non-Identified (Non-ID) processing, and cloud services to make it easier for people to self-respond. This Non-ID methodology has significant impacts on the way USCB staff execute our advertising campaign and promises to increase our self-response rates. The third uses existing administrative data to reduce cost. USCB will use existing data in many components of the 2020 Census including identification of vacant, non-existent, and occupied living quarters for non-respondents. The last area of innovation leverages technology to make field staff more efficient. Much of the field tasks will be automated including recruiting people, hiring staff, training staff, payroll and expense reporting, case assignments, case management, route optimization, quality control, and communications. The 2020 Census promises to be innovative, and improve over past censuses that relied heavily on paper questionnaires.

Keywords: Satellite imagery, internet, cloud services, administrative records, automation

1. Introduction²

In the United States of America (US), Census Bureau staff conduct a census every ten years. The core purpose of the 2020 Census is to count the population and housing, and disseminate the results to the President, the states, and the American people. The census data are used broadly in the US and are summarized in these three main uses of our data:

1. Apportionment of seats allocated to the states for the House of Representatives;
2. Appropriating federal funds to local communities (an estimated \$675 billion annually) [2]; and
3. Establishing the data backbone for the decade to ensure good decision-making and to continue

building confidence in the government, society, and the economy.

This third use of the data remains difficult to track and quantify but many believe this use is the most critical.

Back in 2010 the United States Census Bureau (USCB) planned to increase automation, reduce the use of paper, and become more efficient with mixed results. Census Bureau staff successfully automated the collection of GPS coordinates for all living quarters in the country but the expensive in-person follow-up operation to collect data from non-responders remained a paper-based operation. In the 2020 Census, Census Bureau staff will leverage technology more than ever and our testing showed dramatic improvements in efficiency as a result.

The innovations planned for the 2020 Census are organized into Four Key Innovation Areas:

- Reengineering Address Canvassing: Adding new addresses to the Census Bureau’s address frame using geographic information systems and aerial

¹Disclaimer: This report is released to inform interested parties of research and to encourage discussion of work in progress. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

²Manuscript is an excerpt from 2020 Census Operational Plan V4.0.

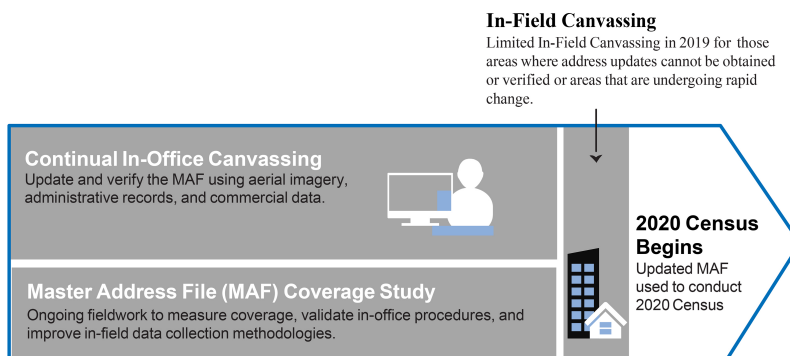


Fig. 1. Summary of reengineering address canvassing.

imagery instead of sending Census Bureau employees to walk and physically check 11 million census blocks.

- Optimizing Self-Response: Encouraging the population to respond to the 2020 Census using the Internet, reducing the need for more expensive paper data capture.
- Utilizing Administrative Records: Using data the public has already provided to the government and data available from commercial sources, allowing realized savings to focus additional visits in areas that have traditionally been hard to enumerate.
- Reengineering Field Operations: Using sophisticated operational control systems to send Census Bureau employees to follow up with nonresponding housing units and to track daily progress.

These four key innovation areas summarize at a high level major innovations planned for the 2020 Census. Many other innovations are planned that are not documented here but can be found in the 2020 Census Operational Plan. Timing and detailed descriptions of all the operations can be found in the 2020 Census Operational Plan, as well. In 2012 Census Bureau staff began a massive research agenda to study the effectiveness of these innovation areas. Over the course of the following 7 years, through 2019, Census Bureau staff conducted over 20 tests to learn, refine, and retest the implementation and application of these operational refinements. The outcome of this research yielded a redesigned 2020 Census with innovation across all the major components of the program.

2. Methodology

The Census Bureau developed innovative methodologies for *Reengineering the Address Canvassing* that

updates the Master Address File/Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) System throughout the decade. Continual research and updating through In-Office Address Canvassing began in September 2015 and continues through 2019 with the establishment of the frame for the 2020 Census. Every year clerks started with the most recent Census Bureau address list and updated it based on new information from the United States Postal Service (USPS) and data from tribal, state, and local governments and third parties (i.e., commercial vendors). Clerks manually compared old aerial (satellite) images to new aerial images in an interactive tool to determine where changes in address inventory occurred, and based on these changes, the Census Bureau captured those updates via In-Office methods. Geographers categorized blocks with changes (active), without changes (passive), and unable to determine (on hold). This plan also included In-Field Address Canvassing for blocks where address updates could not be obtained or verified using In-Office methods (active and on hold). See Fig. 1. The number of addresses requiring In-Field Canvassing is expected to be approximately 35 percent of the total number of addresses.

The goal of *Optimizing Self-Response* is to communicate the importance of the 2020 Census to the U.S. population and generate the largest possible self-response, reducing the need to conduct expensive in-person follow-up with those households. As shown in Fig. 2, the Census Bureau will motivate people to respond by using technology and administrative records and third-party data to target advertisements and tailor contact strategies to different demographic groups and geographic areas. The Census Bureau also will use its partnership program, providing information to government agencies and hosting events at community, recreation, and faith-based organizations. Communication



Fig. 2. Summary of optimizing self-response.

and contact strategies will encourage the use of the Internet as the primary response mode through a sequence of invitations and postcard mailings. In addition, Census Bureau enumerators will leave materials to encourage self-response.

A second key aspect of Optimizing Self-Response is to make it easy for people to respond from any location at any time. This is done in several ways:

- By enabling people to respond via multiple modes (Internet, paper, or telephone if they call the Census Questionnaire Assistance Center).
- By allowing respondents to submit a questionnaire without a unique identification code (Non-ID Response).
- By providing online forms in multiple languages.

For these innovations to be successful, respondents must know that their personal information is protected. Thus, a key element of this innovation area is to assure respondents that their data are secure and treated as confidential. The basic premise of Non-ID processing starts with respondents giving us their address that Census Bureau staff use to get their census ID off the Master Address File for them. For cases that fail ID address matching, automated geocoding occurs, followed by a clerical matching and geocoding step, if needed. This Non-ID processing allows respondents to self-respond anytime and anywhere, including but not limited to sporting events, waiting for the bus, or while watching television. Both ID and Non-ID Internet Self-Response will be hosted in the cloud. This IT infrastructure allows for basically unlimited numbers of users without performance degradation. The USCB is committed to performance testing these systems and verifying this critical scaling stability. Another key area of effort at the USCB revolves around

stability of our Internet Self-Response systems from outsider threats. With secure systems that scale to extremely large volumes, the USCB remains committed to these innovations.

The goal of *Utilizing Administrative Records* is to use information people have already provided to improve the efficiency and effectiveness of the 2020 Census, and in particular, reduce expensive in-person follow-up activities. Administrative record data refers to information from federal and state governments. Third-party data refers to information from commercial sources like credit brokers. As shown in Fig. 3, data from both sources help to improve the quality of the address list (frame), increase the effectiveness of advertising and contact strategies, validate respondent submissions, and reduce field workload for follow-up activities. As has been done in prior decades, administrative data from the USPS and other government records are used to update the address frame and reflect changes to the housing stock that occur over time. Additional administrative records sources, as well as third-party data from commercial companies, will also be used for this purpose. In addition, these data sources will be used to validate incoming data from tribal, federal, state, and local governments.

To increase the effectiveness of advertising and contact strategies, the Census Bureau will use demographic and geographic information from various administrative record and third-party data sources to help target the advertising to specific populations.

Administrative records and third-party data will also be used to validate respondent addresses for all Internet responses. This will help prevent fraudulent and erroneous submissions.

Finally, a primary use of administrative records is to reduce field workload for follow-up activities. To

Improve the quality of the address list.	Update the address list.	Validate incoming data from tribal, federal, state, and local governments.
Increase effectiveness of advertising and contact strategies.	Support the micro-targeted advertising campaign.	
Validate respondent submissions.	Validate respondent addresses and prevent fraudulent submissions.	
Reduce field workload for follow-up activities.	Remove vacant, nonexistent, and nonresponding occupied housing units from the NRFU workload.	Optimize the number of contact attempts. Tailor work assignments with best time of day to contact.

Fig. 3. Summary of utilizing administrative records and third-party data.

this end, the Census Bureau will use data from internal and external sources, such as the 2010 Census, the USPS, the Internal Revenue Service, the Social Security Administration, the Centers for Medicare and Medicaid Services, and the American Community Survey to identify when nonresponding housing units should be classified as occupied and that high-quality administrative data can be used for the enumeration. These units will be visited one time in person during non-response follow-up and, if not enumerated during that visit, after seven days they will be mailed a postcard encouraging self-response and removed from the non-response follow-up workload for all subsequent activity. Addresses found to be “administrative records vacant” or “administrative records nonexistent” will be mailed a postcard first and then they will be visited once during the non-response follow-up before they are removed from the non-response follow up workload. Data from these sources will also be used to tailor work assignments related to the best time of day to contact a household.

The goal of the *Reengineering Field Operations* innovation area is to use technology to efficiently and effectively manage the 2020 Census fieldwork, and as a result, reduce the staffing, infrastructure, and brick-and-mortar footprint required for the 2020 Census. Figure 4 shows the main components of the reengineered field operations: increased use of technology and increased management and staff productivity.

The 2020 Census field operations will rely heavily on automation. For example, the Census Bureau plans to provide most listers and enumerators with the capability to work completely remotely and perform all administrative and data collection tasks directly from a mobile device. Supervisors will also be able to work re-

motely from the field and communicate with their staff via these devices. These enhanced capabilities significantly reduce the number of offices required to support 2020 Census fieldwork. In the 2010 Census, the Census Bureau established 12 Regional Census Centers (RCCs) and nearly 500 Area Census Offices (ACOs). In 2010 the agency hired more than 516,000 enumerators to conduct Nonresponse Followup (NRFU) activities. The new design for the 2020 Census field operations includes six RCCs with 248 ACOs to support 350,000 enumerators.

In addition, automation enables significant changes to how cases are assigned and the supervision of field staff. By making it easier for supervisors to monitor and manage their workers, the ratio of workers to supervisor can be increased, reducing the number of supervisors required. This stream-lines the staffing structure. Other design changes include optimized case assignment to alter the days of the week and times of the day to visit addresses and optimizing routing staff to those assignments to travel shorter distances and reduce travel time.

All administrative functions associated with most field staff will be automated, including recruiting, hiring, training, time and attendance, and payroll. Finally, the new capabilities allow for quality to be infused into the process through alerts to supervisors when there is an anomaly in an enumerator’s performance (e.g., the Global Positioning Satellite indicator on an enumerator’s handheld device indicates that she or he is not near the assigned location) and real-time edits on data collection. Accordingly, the quality assurance process used in the 2020 Census is being improved to account for changes in technology.

Increased Use of Technology	Increased Management and Staff Productivity
<ul style="list-style-type: none"> • Automated and optimized work assignments • Automated recruiting, training, payroll, and expense reporting • Automated applications for address canvassing and enumeration on mobile devices 	<ul style="list-style-type: none"> • Increased visibility into case status for improved workforce management • Redesigned quality assurance operations • Improved communications

Fig. 4. Summary of reengineering field operations.

3. Results

The 2020 Census promises to be much more efficient than prior censuses. The In-Field Address Canvassing operation has been reduced from 99 percent in field during the 2010 Census down to about 35 percent projected for the 2020 Census. This reduction in fieldwork decreases the field costs from \$443 million [4] in 2010 to about \$100 million expected for 2020. For self-response, by making it easier for people to respond through Non-ID processing and offering an Internet option, our response rates are expected to remain stable. Although this may seem like steady efficiency, self-response rates are consistently declining in the US across government and private surveys. By utilizing administrative records data, Census Bureau staff plan to reduce Nonresponse Followup workloads by about 20 percent. Recognizing the estimated cost of our most costly field operation (Nonresponse Followup) at about \$1.4 billion [1], cost avoidance continues for this innovation area. Finally, the field workers look to increase efficiency by about 50 percent for the fieldwork. The efficiencies expected for the recruiting, hiring, training, payroll, case assignment, case management, routing, quality control, and communication contribute to major cost savings. By moving from paper-based processes in 2010 to a fully automated process in 2020 the efficiency of our personnel practices is expected to dramatically improve.

4. Discussion and conclusion

The 2010 Census cost the citizens of the United States just over \$10.2 billion in nominal dollars and \$12.1 billion adjusted for 2020 Census inflation and population/housing unit growth. When the 2020 Census program started in 2012, one goal was to maintain costs compared to 2010 adjusted for inflation and population/housing unit growth. The current estimates for

the 2020 Census are approximately \$12.7 billion [1].³ Although Census Bureau staff estimated our IT costs for the 2020 Census at over \$5 billion, the innovations are surely helping to avoid costs in other areas. The innovations proposed, researched, and implemented over the last decade span the entire survey lifecycle. Major innovations on:

1. how Census Bureau staff maintain our address list,
2. how Census Bureau staff encourage our people to self-respond,
3. how Census Bureau staff leverage existing data, and
4. how Census Bureau staff manage our massive field staff

promise to make the 2020 Census much more efficient and cost effective. As the 2020 Census approaches, this redesign of our methodology brings excitement of innovation to our staff, our stakeholders, and the American people.

References

- [1] Budget Staff. (2019). *2020 Census Life-Cycle Cost Estimate Executive Summary* (pg 25). Washington, DC: U.S. Census Bureau.
- [2] Hotchkiss M. A. (2017 September). *Uses of Census Bureau Data in Federal Funds Distribution*. Retrieved from U.S. Census Bureau, Washington, DC: www.census.gov/library/working-papers/2017/decennial/census-data-federal-funds.html.
- [3] Operational Plan Team. (December 2018). *2020 Census Operational Plan V4.0*. Washington, DC.
- [4] Team A. L. (2012). *2010 Census Address Canvassing Operational Assessment*. Washington, DC: 2010 Census Planning Memoranda Series No. 168.

³This is a point estimate for the programmatic components of the 2020 Census. The official estimate includes component of risk and uncertainty, which could bring the cost of the 2020 Census to \$15.6 billion.