Daniel Castro

Senior Analyst
Information Technology and Innovation Foundation (ITIF)

"Census: Learning Lessons from 2010, Planning for 2020"

Before the

Senate Homeland Security and Government Affairs Committee
Subcommittee on Federal Financial Management, Government Information, Federal Services, and International Security
U.S. Senate

April 6, 2011

Chairman Carper, Ranking Member Brown and members of the Committee, I appreciate the opportunity to appear before you to discuss technology opportunities and trends for the 2020 Census. My name is Daniel Castro. I am a senior analyst at the Information Technology and Innovation Foundation (ITIF). ITIF is a nonpartisan research and educational institute whose mission is to formulate and promote public policies to advance technological innovation and productivity.

The Constitutionally-mandated decennial census produces important demographic data used by government to deliver services and make political, economic and social decisions. Enumerating the entire population of our nation is a considerable undertaking and, as far back as the 1890s with the invention of Herman Hollerith's electro-mechanical tabulator, the government has relied on information technology (IT) to make the process more efficient. In 2000, the U.S. Census Bureau established itself as a digital pioneer by making the United States one of the first countries in the world to use an Internet-based data collection method for its population census.

Through this project, the United States demonstrated that it was a global leader in using IT for e-government.

Unfortunately, the Census Bureau has failed to keep pace with current advancements in IT. The 2010 Census cost approximately $13 billion, more than any other census before it, and did not use IT efficiently or cost-effectively.[1] For example, the Census Bureau did not provide an online response option and its efforts to use mobile devices for non-response follow-up were unsuccessful. While we cannot roll back the clock, we can learn from the past, as well as the successes of other nations, and better use IT in the next census. With that in mind, to improve the 2020 Census, I recommend the Census Bureau provide an online data collection option and adapt to current technology trends such as cloud computing, low-cost mobile devices, and social networks.

## Provide an Online Data Collection Option for the 2020 Census

The Census Bureau should allow individuals to submit their census forms online in 2020. Internet-based transactions are generally less costly, more accurate and more secure than their paper-based counterparts. The Census Bureau itself experienced generally positive results in its earlier tests of an Internet response option in 2000, yet it failed to fully explore how most effectively to implement a similar approach in the 2010 Census. During this period, home broadband adoption has grown in the United States from 3 percent in June 2000 to 66 percent in May 2010.[2] Moreover, by 2020 the U.S. Federal Communications Commission estimates that more than 90 percent of American home will have broadband access.[3] It is a positive step that the Census Bureau is currently testing an internet response option for the American Community Survey to better plan for the 2020 Census.[4]

My review of the census data collection methodologies used by government statistics agencies in countries other than the United States shows that other countries, including Canada, Singapore, Norway, and Australia, have far surpassed the United States in the use of the Internet to conduct the census. For example, the Australian Bureau of Statistics expects that at least 30 percent of the population will choose to use the Internet response option in 2011, almost tripling the percentage of households that completed it in 2006.[5] Worldwide more than 30 countries are providing or experimenting with an Internet response option from Japan to Slovakia.[6]

Many countries have achieved notable levels of success with an Internet response option. In 2006, Canada found that if it achieved a 16 percent Internet response rate, the cost savings would exceed the cost of implementing the Internet response option. Given that it exceeded its response rate goal, Canada not only saved money on its 2006 census but now has the infrastructure in place to conduct future censuses online.[7] For example, for the 2011 Census Statistics Canada announced it was spending only $19.7 million (CAD) for an upgrade of the software used in the 2006 Census.[8]

Moreover, developing an online census, or any similar IT application, has high fixed costs and low marginal costs. This means that while the development costs for most countries are similar, large countries can spread the costs over a larger population, resulting in a lower cost per person. Given that the population of the United States is almost 10 times larger than the population of Canada, the United States should expect to be able to develop a cost-neutral Internet response option even with a significantly smaller percentage of its population using the option.

Singapore also generated an impressive Internet response rate of 15 percent in its 2000 census, even though only 32 percent of Singapore's population had Internet access at the time.[9] Last

year, Singapore conducted its 2010 Census and the online response rate grew to 38 percent.[10] This high response rate has been credited in part to the design of Singapore's census, which encouraged respondents to use the Internet. The Singapore census offered three data response methods—mail, telephone, and Internet—but each response method was offered at a different stage. First, all individuals were invited to respond by Internet. Then for the individuals who had not responded by Internet, phone interviews were offered. Finally, census workers used paper forms to enumerate the remaining individuals.[11]

Providing the U.S. population with an Internet-based census response option would yield many benefits for both the Census Bureau and respondents, including increased convenience, accessibility and usability, improved accuracy, reduced costs, increased security over traditional paper-based methods, and other less tangible benefits.

Web-based forms can provide a more user-friendly experience than paper forms by offering well-designed interfaces that increase legibility and accessibility. For example, people with disabilities may find an online form easier to complete and return than a paper-form because of the accessibility features available on modern computers, including assistive devices, large text, and screen readers. In addition, web-based forms can increase usability by providing additional guidance and contextual help. To illustrate, web surveys can use popup windows, trouble-shooting wizards, and searchable guides to ensure individuals can get help and successfully complete the questionnaire. Furthermore, the web-based form can be customized for each user. Thus, for instance, questions can be personalized on the basis of previously provided information, and irrelevant questions can be suppressed. Users can also select their preferred language. Finally, users can easily print a copy of the census questionnaire for their own records after completing the form.[12]

Collecting data online has the potential to improve data accuracy over paper-based methods for several reasons. First, online forms can handle atypical data such as unusually long responses. One limitation of the paper-based census, for example, is that some large households cannot report all information on each household member because there is not space on the questionnaire. With a web-based census form, this space problem would not exist. Second, web forms can have automatic error-checking to prevent users from submitting invalid or incomplete data. Thus, for example, client-side scripts—computer code executed in the user's browser—can generate a warning to alert users that they have not completed all fields before they submit their web form. Third, Internet-based census forms can use discrete options, such as selection menus or checkboxes, to reduce the number of possible ambiguous answers and ensure all data is coded correctly.[13] Fourth, online data collection eliminates the possibility for error during the scanning, decoding, and transcribing processes that must occur for paper forms. In addition, because data collected online require less processing than data collected using paper forms, census results can be made available more quickly if the data are collected online.

Using the Internet to collect census data can help reduce the costs of data collection by reducing mail-back costs, processing costs, and follow-up costs.[14] First, the Census Bureau saves the cost of postage that it must pay for every census questionnaire returned by mail. Second, Internet-based web forms eliminate the expensive intermediary steps of paper-based data collection, such as scanning and transcribing. Third, providing an Internet response can help the Census Bureau reduce the number of individuals with whom it must follow up. Some of the savings could come from individuals who respond to the census by Internet who otherwise would not have responded, although previous surveys have found that offering an Internet response option did not increase the overall response rate. Additional savings could come from reduced follow-up

costs, because forms submitted online are less likely to be inconsistent or incomplete.[15] Further savings could come from reducing nonresponse follow-up costs for replacement questionnaires, which are sent to individuals who have not responded after some predetermined threshold date. The Census Bureau found during tests in 2003 and 2005 that when an Internet census response option is provided, the response rate prior to the threshold date is higher. The Census Bureau has estimated these savings from reduced nonresponse follow-up to be in the range of $300,000 to $500,000 per census.[16]

In terms of data security, it should be noted that all census records are eventually stored electronically, regardless of how they are collected. The security of the census data while in storage is not affected by the manner in which the census data are collected. Internet-based data collection affects the security of the data only while the data are in transit. If strong encryption and authentication methods are used, sending data over the Internet can be more secure than sending a census questionnaire through the mail. This statement is not an indictment of the U.S. Postal Service but rather a recognition of the fact that encrypted digital information is more secure than the same unencrypted information when it is on paper. Encrypted electronic data, for example, can only be accessed by an individual with the appropriate passcode, whereas unencrypted data in an envelope can be accessed by anyone with a letter-opener.

Finally, providing an Internet response option for the census has a number of additional, less quantifiable benefits. Providing respondents with access to the census questions online can help the Census Bureau create a more modern image for itself and might help reduce some of the resistance to the census.[17] Citizens increasingly expect government services to be online, and providing an Internet option would help the Census Bureau meet the demands and expectations of its customers.[18] Large-scale projects such as providing an Internet response option for the

census could also help to develop the skills and knowledge base of federal employees, which could then subsequently be leveraged in other e-government projects.

## Use Technologies Such As Cloud Computing, Mobile Devices and Social Networks

Since planning for the decennial census begins well in advance of the actual operation, the Census Bureau cannot be expected to predict the technology that will be available many years down the road. However, the Census Bureau can incorporate technology trends into its planning and operations. For example, while we did not know in 2000 how many Americans would be using the Internet a decade later, we had every reason to believe that this upward trajectory in adoption would continue. With this in mind, the Census Bureau should plan to make better use of technology such as cloud computing, mobile devices, and social networks—three trends in IT which promise to deliver flexibility, cost-savings, and user-engagement to organizations that employ them wisely.

The first IT trend that the Census Bureau should incorporate in to its 2020 Census is cloud computing. Cloud computing is a term that refers to the practice of selling IT as a service, including software as a service, platform as a service, and infrastructure as a service.

In the old model of computing, an organization would estimate how much computing power it needed, and then purchase the number of servers required to meet its peak needs. Most of the time, however, these computing resources would be underutilized. In addition, if an organization's needs exceeded it estimates, the organization would have to scramble to purchase and bring online more servers. Cloud computing eliminates many of these challenges.

For one, cloud computing creates a more flexible environment that allows organizations to "rent" computing power on an as-needed basis—an organization can scale up or down its IT usage according to demand. Organizations also benefit from the agility that cloud computing offers them as they have no long-term commitments and no high-fixed costs. Government agencies, for example, can better align cost with use by only paying for their actual use of IT resources, rather than having to overbuild capacity based on potential demand. This agility also allows organizations to easily upgrade their applications as they can change platforms simply by switching cloud providers.

Cloud computing also allows organizations to focus on their core business and not their IT. Running data centers—buying, installing, operating, maintaining, and upgrading servers—is resource intensive. Organizations benefit from cloud computing because cloud computing service providers can provide greater economies of scale, share resources across multiple customers, and provide higher levels of expertise in operating a secure, reliable, and energy efficient data center. Cloud computing can be more efficient as most organizations, particularly small and mid-sized organizations, are not sufficiently large or sophisticated enough to take advantage of the economies of scale possible with IT.

The concepts behind cloud computing—on-demand, scalable and pay-per-use—make it ideal for applications such as the census which have variable demand for resources. For example, the computing resources needed by the Census Bureau peaks sharply during the rather short period of time when individuals and census workers are submitting responses, but go unused at other times. This means that if the Census Bureau, or its contractor, used cloud computing, it would not need to invest in a large amount of IT infrastructure but could instead only pay for the actual computing resources used. A move to the cloud for the census would also be in line with the

Office of Management and Budget's (OMBs) efforts to create a "cloud first" policy that is reducing the size of the federal government's data center footprint.[19]

Any technology investment carries risks, but most of these risks can be managed through proper administrative and technical controls. Unfortunately, government IT investments are often judged primarily on their performance rather than both their performance and cost-effectiveness. When performance is the dominant metric, government agencies have a tendency to delay adopting new technology that may introduce some risk. For example, the Census Bureau might be tempted to eschew a public cloud offering (i.e. a cloud service provided in a shared environment over the Internet) in favor of a more expensive private cloud solution (i.e. a cloud service deployed on a private network that shares no resources) because of potential security or privacy concerns or fears that a public cloud may not offer sufficient reliability.[1] To avoid this type of risk-averse decision-making that doomed the use of an Internet response option in 2010, the Census Bureau should be held accountable not only for its performance but the degree to which it applies cost-saving technologies.

The second IT trend that the Census Bureau should take into account is the proliferation of low-cost mobile devices, such as smartphones and tablet PCs, and the mobile Internet. By 2015, fourth-generation (4G) wireless networks such as LTE should be widely deployed and available to provide mobile broadband access to over 90 percent of Americans. Moreover, we should expect this trend of increased mobile connectivity to continue.

[1] In regards to reliability, cloud computing is somewhat analogous to electricity. Just as government could have more reliable operations if it ran its own generators rather than rely on public utilities, the benefits of doing so do not outweigh the costs. Similarly, government agencies should be cautious that they do not avoid public cloud offerings simply because of nebulous concerns about data security.

The Census Bureau should plan to take advantage of the availability of low-cost, high-performance, off-the-shelf mobile computing devices in 2020. Using a mobile device for data collection and address canvassing can allow census workers to enter data more accurately and efficiently. Over 100 hundred countries have reported using GPS/GIS technology for automating geographic data collection and verification in their census.[20] The Census Bureau planned to use handheld computers in 2010, but ultimately abandoned this effort due to mounting delays and costs. Rather than developing proprietary and expensive handheld devices as the Census Bureau choose to do for 2010, in the future it should use low-cost off-the-shelf equipment. This will allow it to select the most cost-effective technology available, rather than using more expensive technology that is almost a decade old.

Because web applications can be developed to be platform-neutral, the Census Bureau does not need to know the exact technical specifications of the next generation of mobile devices. It does not matter whether these devices are smartphones, PDAs, tablets or some hybrid form-factor as long as they can access the Internet. By using light-weight mobile apps that run from the cloud, the Census Bureau can develop data collection tools for census workers that will run on tomorrow's devices. In addition, if the Census Bureau uses off-the-shelf products, like a consumer-grade tablet PC, in 2020, it could then donate these computers to low-income schools after the census is complete. Donating these PCs would not pose any privacy risks because these computers would have no sensitive data stored on them.

Consider the example set by Brazil, which is the fifth largest country in the world by population and has a population that is roughly two-thirds that of the United States. Brazil equipped its enumerators in the 2010 Census with off-the-shelf equipment—150,000 LG smartphones and 70,000 Mio PDAs—for its data collection and mapping efforts.[21] In part because of the nation's

strategic use of off-the-shelf technology, Brazil conducted its 2010 Census for approximately $900 million.

Finally, the Census Bureau should more actively engage with users on mobile devices and social networks. The purpose of these activities should be to continue what the Census Bureau has always done which is to bring the census to the people. Achieving high response rates is important for reducing the need for non-response follow-up, one of the most costly aspects of the census. One way to improve response rates is to engage with users online. For example, individuals that complete the census online could be given the chance to share a "digital badge" with their contacts on social networks that says they have been counted which would serve as a reminder to others to complete their forms. Currently, 46 percent of U.S. adults use social networks.[22] Of U.S. adult Internet users, 61 percent use an online social network—including more than 40 percent of U.S. Internet users use Facebook and 8 percent are on Twitter.[23] The use of social networks is expected to continue to continue to grow over time.

Increasingly Americans will also be accessing the Internet through mobile devices. The Census Bureau should be sure to consider the different types of devices someone might use to respond to an online questionnaire and incorporate tools to make it easier for individuals to complete a census form using a mobile device. Using technology like QR codes (i.e. matrix barcodes), for example, could give individuals the ability to point the camera of their smartphone or tablet at a census form and be automatically directed to their census form online without keying in an alphanumeric passcode.

## Conclusion

With strong program management and proper IT governance, the Census Bureau should be able to better utilize IT in the 2020 Census to not only to improve existing operations, but to find innovative ways to use technology to deliver more value to citizens. For example, the Census Bureau (or Congress) may eventually decide that collecting data every decade no longer makes sense in our fast-paced world and instead turn to population registers for this data. Many countries, including Sweden, Finland, Denmark, and Iceland, for example, no longer conduct a questionnaire-style census but instead gather census data from administrative records. These countries have found that a register-based census—that is, a census using existing data from administrative records—is cheaper and faster and places less of a burden on the county's population than a questionnaire-style census.[24] In addition, some countries have begun substituting or supplementing a traditional population-wide census conducted every five or 10 years with smaller surveys conducted more regularly.

Given the rising costs of conducting the decennial census, the U.S. Census Bureau should welcome the opportunity to use IT to reduce costs and improve quality. Every technology project faces certain obstacles; however, the Census Bureau should respond to these challenges with creativity and thoughtfulness, rather than resign itself to the familiar yet impractical methods of the past. Certainly new technology is not a panacea, but it can help organizations like the Census Bureau achieve their mission more efficiently and effectively.

## Endnotes


1. "2010 Census: Data Collection Operations Were Generally Completed as Planned, but Long-standing Challenges Suggest Need for Fundamental Reforms," Government Accountability Office, December 2010, http://www.gao.gov/new.items/d11193.pdf.

2. "Broadband and Dial-up Adoption, 2000-1010," Pew Internet & America Life Project surveys 2000-2010, http://www.pewinternet.org/Trend-Data/Home-Broadband-Adoption.aspx (accessed April 1, 2011).

3. "National Broadband Plan: Connecting America," Federal Communications Commission, http://www.broadband.gov/plan/2-goals-for-a-high-performance-america/.

4. Todd Hughes, "The American Community Survey," International Census Forum, October 12, 2010, http://www.census.gov/icf/docs/acs.pdf.

5. "2011 Census Fact Sheet", Australian Bureau of Statistics, March 15, 2011, http://www.abs.gov.au/websitedbs/D3310114.nsf/home/2011+Census+Fact+Sheet:+eCensus.

6. Yacob Zewoldi, "Seminar on New Technologies in Population and Housing Censuses: Country experiences" United Nations Statistics Division, Side event of the 42nd session of the, Statistical Commission, February 24 2011, New York.

7. Statistics Canada, "Census Technology: Recent Developments and Implications on Census Methodology," presented to the Group of Experts on Population and Housing Censuses, Conference of European Statisticians, Economic Commission for Europe, Tenth Session, Astana, Kazakhstan, 4-6 June 2007, http://www.unece.org/stats/documents/ece/ces/ge.41/2007/9.e.pdf.

8. "Private sector involvement," Statistics Canada, November 12, 2010, http://census2011.gc.ca/ccr01/ccr01_009-eng.htm.

9. Note that the census only surveyed 20 percent of the population. Source for Internet users in Singapore: "2000: Singapore," ICT-Eye Database, International Telecommunications Union (ITU), accessed on December 1, 2007, http://www.itu.int.

10. Chan Herng Wei, "Census of Population 2010 – Increased Use of Internet in Census Submission," Statistics Singapore Newsletter, March 2011, http://www.singstat.gov.sg/pubn/papers/people/ssnmar11-pg1-7.pdf.

11. Werner Haug, Swiss Federal Statistical Office, "Population Censuses on the Internet," IUSSP General Population Conference 2001, Salvador de Bahia, Brazil, 18-24 August 2001, http://iussp.org/Brazil2001/s00/S06_01_Haug.pdf.

12. Ibid.

13. Ibid.

14. Ibid.

15. Deborah H. Griffin, Donald P. Fischer, and Michael T. Morgan, "Testing an Internet Response Option for the American Community Survey," presented at the annual conference of the American Association for Public Opinion Research, Montreal, Quebec, May 2001, downloaded from the U.S. Census Bureau, http://www.census.gov/acs/www/Downloads/ACS/Paper29.pdf.

16. MITRE Corporation, "An Assessment of the Risks, Costs, & Benefits of Including the Internet as a Response Option in the 2010 Decennial Census, Version 2.0," prepared for the U.S. Census Bureau (McLean, VA: MITRE Corporation, January 17, 2007).

17. Haug Wei, "Census of Population 2010 – Increased Use of Internet in Census Submission."

18. Anil Arora and Graeme Gilmou, Statistics Canada, "Statistics Canada—Census on the Net," presented at Conference of European Statisticians, United Nations Statistical Commission and United Nations Economic Commission for Europe, Geneva, Switzerland, 13-15 June 2005, http://unece.org/stats/documents/ces/2005/11.e.pdf.

19. Jeffrey Zients, "Driving IT Reform: An Update," Office of Management and Budget, November 19, 2010, http://www.whitehouse.gov/blog/2010/11/19/driving-it-reform-update.

20. Yacob Zewoldi, "Introduction: Seminar on New Technologies in Population and Housing Censuses: Country experiences," United Nations Statistics Division, February 24, 2011, http://unstats.un.org/unsd/statcom/statcom_2011/Seminars/NewTechnologies/UNSD%20-%20Seminar%20on%20New%20Technologies%20f.ppt.

21. Eduardo Nunes, "IT in the Brazilian 2010 Census," Instituto Brasileiro de Geografia e Estatística, February 24, 2011, http://unstats.un.org/unsd/statcom/statcom_2011/Seminars/NewTechnologies/IBGE%20-%20Brazilian%20Census%202010%20_%20UNSC%20_%2021%20Feb%202011.pptx.

22. "Global Publics Embrace Social Networking," Pew Research Center, December 15, 2010, http://pewglobal.org/2010/12/15/global-publics-embrace-social-networking/.

23. "Trend Data," Pew Internet & American Life Project, September 2010, http://www.pewinternet.org/Static-Pages/Trend-Data/Online-Activites-Total.aspx, "United States of America Internet and Facebook Users Stats," Internet World Stats, August 31, 2010, http://www.internetworldstats.com/stats26.htm, and Aaron Smith and Lee Rainie, "8% of online Americans use Twitter," Pew Internet & American Life Project, December 9, 2010, http://www.pewinternet.org/Reports/2010/Twitter-Update-2010.aspx.

24. Åke Bruhn, "The 2005 Population and Housing Census in Sweden Will Be Totally Register-Based," Symposium on Global Review of 2000 Round of Population and Housing Censuses: Mid-Decade Assessment and Future Prospects, 6 July 2001 Statistics Division, Department of Economic and Social Affairs, United Nations Secretariat, New York, New York, 7-10 August 2001, http://unstats.un.org/unsd/demographic/meetings/egm/Symposium2001/docs/symposium_13.htm.