

Envisioning Census 2030: A 21st Century Census Curated Data Enterprise

April 2021

Sallie Keller, Distinguished Professor in Biocomplexity, Director of the Social & Decision Analytics Division, Biocomplexity Institute at the University of Virginia, sak9tr@virginia.edu

Ken Prewitt*, Carnegie Professor of Public Affairs and Special Advisor to the President, Columbia University, U.S. Census Bureau Director, 1998-2001, kp2058@columbia.edu

*This initiative was co-developed by Sallie Keller and Ken Prewitt with extensive input from experienced census experts. Just recently, Prewitt shifted from his direct involvement of this work to take on a new role as special advisor to the U.S. Census Bureau Director.

Purpose Statement

We propose a set of revolutionary advancements to help the U.S. Census Bureau (the Bureau) implement a new approach for the 2030 Decennial Census. To better meet the country's needs, we offer innovations whereby essential data will be built over the decade with the goal of providing an even more complete and accurate enumeration of the United States population in 2030. Along the way, the Bureau will innovate its operations and produce more timely, more robust, and more accurate findings to serve the nation's need for data about its people, places, and economy and to more fully reflect the diversity of the nation's racial and ethnic composition. The Bureau will achieve these advancements while reducing public burden, leveling out the cost of its enumeration operations across the decade, better serving more geographic entities, and maintaining the essential trust that the collected data are secure.

Intent of the Initiative

Our nation's Framers placed a decennial census in the U.S. Constitution to ensure the fair redistribution of political representation as the new nation grew and evolved, commencing a national expectation for precision and quality that persists to this moment.

Today, the nation relies on the flow of relevant, accurate, timely, reliable, and objective statistics to support the decisions of governments, businesses, individuals, households, and other organizations. Federal statistical agencies release many of the statistics available about the United States' economy, population, natural resources, environment, and public and private institutions. It is the responsibility of Federal agencies engaging in statistical work to support the quality and accessibility of the Federal statistical information our nation uses to monitor and assess performance, progress, and needs.

The Bureau's mission is to serve as the nation's leading provider of comprehensive, quality data about its people, places, and economy. Census data are a gold standard, foundational to all official statistics. They are a key public good accessible to all, and which support federal, state, and local planning, as well as economic activity.

To be useful, these data need to be available for all areas and subgroups of the nation. They must serve tiny communities in rural areas to discrete urban neighborhoods while accurately reflecting the full diversity of the nation's racial and ethnic composition.

The Bureau is experiencing increased demand for more data and more timely data delivery, while confronting challenges to traditional data collection methods. The Covid-19 pandemic accelerated the Bureau's pace of innovation through adaptation to relentless forces that are changing the shape of data collection in official surveys.

New approaches are required to help reduce the historical differential undercounts of race and ethnic subgroups, and young children. Public trust in the new national commitment to evidenced-based policy making starts with confidence that the underlying data accurately and equitably reflect the full diversity of the nation.

Looking ahead, the Bureau is moving to increase operational efficiencies, exploring how to harvest new data sources, building data innovations like the Household Pulse Surveys, deploying new statistical methods, and working to improve data quality and availability. These changes are required to meet the needs of more stakeholders and users with greater analytical capabilities and expectations.

Central to these innovations is researching how to expand the use of public and private administrative data, other survey data, and third-party data. These include commercial and transactional sources that capture the same content currently collected by the Bureau or provide new contextual information in combination with Census data. Exploiting these additional sources can help deliver official data faster, enhance efficiency, improve accuracy while reducing respondent burden.

Equally central to the innovation are operational reforms designed to integrate or connect data across administrative record, survey, and other data the Bureau possesses into a curated data enterprise linking disconnected survey frames on businesses, housing units, jobs, and people into a unified resource. The Bureau's universal frame is intended to bring together the Business Register of establishments, the Master Address File of housing units, the Longitudinal Employer-Household Dynamics (LEHD) jobs frame, and a demographic frame of individuals geo-coded to a multitude of larger and smaller geographic areas.

Introducing the 21st Century Census Curated Data Enterprise

An integrated universal statistical frame will form the foundation of a new data infrastructure we describe as the 21st Century Census Curated Data Enterprise (CDE). The 21st century model we have in mind involves multiple-sourced data products that, taken together, can build a new, more accessible, and reliable information platform for the nation.

Expanding the use of all data sources and integrating across the survey frames should provide more timely and more robust data for smaller geographies, rural areas, Tribal lands, and subpopulation groups that are less well served by current Bureau data products, including decennial census data files.

Linking across time, and surveys, supplemented with third-party and administrative data can help improve the coverage and completeness of Census data, innovating beyond a survey-based model that has reached scientific and practical limitations. The CDE's multi-source data innovation offers a better option to help reduce differential undercounts of children as well as race and ethnic populations than a survey or census enumeration alone.

The CDE can be augmented over the decade with multi-source government and private sector data sources. Starting with the 2020 decennial census records, we can populate the curated platform with rich social and economic data at a scale the country has not previously imagined across a wider array of geographic levels.

This modernization, founded on the 2020 decennial and across all the Bureau's surveys and censuses, can transform the manner in which the Bureau conducts the 2030 decennial. It can lessen respondent burden, spread the enumeration life-cycle cost more evenly across the decade, and reduce the massive scale of decennial operations while improving coverage and accuracy. The 2030 Census will be the beneficiary of tested innovations on data collection and integration in the early part of the decade with better outcomes than could be achieved without these advancements.

The CDE will introduce greater equity in the decennial operational design. The new design should be less sensitive to the variations in disparate census promotional and outreach efforts intended to garner high levels of self-response. At the same time, it will give advance notice about geographies stakeholders and the Bureau together can target to resolve the most vexing non-response challenges.

Along the way, the CDE will have the added benefit of allowing researchers to query databases instead of having to link data across siloed frames, producing more robust outcomes.

Researchers will be able to link people to jobs to businesses, making new analyses possible, and the ability to study, for example, the demographic characteristics of workers in a business or industry or to answer national policy questions at relevant geographic specificities such as whether Opportunity Zone companies are hiring people who live in that zone.

We acknowledge these innovations require greater attention to protecting the security, privacy, and confidentiality of all data. Census has always had a legal, statistical, and moral obligation to protect the confidentiality of the information it collects. Respondent participation increases when people believe their data will remain confidential.

The decennial is fundamental to public policy goals that seek to achieve racial and ethnic equity. To meet these needs and adapt to changing perceptions of identity, innovations in collecting and measuring these characteristics should seek to improve upon existing data standards.