

RESPONSE TO NAC ARTPD WG RECOMMENDATIONS

The Working Group on Administrative Records and Third Party Data (ARTPD) assessed the use of external data sets to improve the 2020 Census. The Working Group explored privacy and data quality issues through directed readings and interviews with subject matter experts in industry, government, and academia. The Census Bureau staff who facilitated the Working Group agrees that external sources should be integrated when the data is accurate and appropriate data stewardship practices are observed. The Census Bureau is committed to ongoing research to improve linkages and coverage of hard to reach populations.

Responses to Recommendations:

1) Data sources should be distinguished as either Administrative Records (AR) or Third Party Data (TPD) in Census Bureau polling, communications, reports, etc.

The Census Bureau will advise users to distinguish between AR and TPD when describing their research and operations. When AR and TPD Information Owners approve uses of their data, they will inform project contacts on appropriate terminology.

2) Any ARTPD under consideration by Census Bureau for decennial census should assess privacy and inclusiveness issues.

The Census Bureau continues to monitor and assess the implications of using external data sources. In FY2015, the 2020 Census Privacy and Confidentiality Team will resume activities. Inclusiveness issues in ARTPD are being explored by the 2020 Census Fitness for Use Project in FY2014 and FY2015. They will be evaluating the coverage and quality of ARTPD across demographic characteristics and file sources.

3) Census Bureau should consider public perception of ARTPD sources in making its decision about how to use ARTPD and develop public messaging strategies on the use of ARTPD.

The Census Bureau is committed to providing quality data on the nation's people and economy while honoring privacy and protecting confidentiality. Public perception is one of the foremost issues concerning use of ARTPD the Census Bureau strives to examine and address the public perception issues surrounding ARTPD use. The Research & Methodology, Decennial, and Communications Directorates are developing messaging strategies and communications plans regarding ARTPD.

4) TPD containing Personally Identifiable Information (PII) should not be used if it was obtained surreptitiously or by "scraping" the web and there should be a presumption against the use of TPD containing PII. That presumption can only be overcome by the Census Bureau staff advocating for its use, when a rigorous analysis similar to that outlined in the attached "decision" tree, clearly and unambiguously shows that the data cannot be obtained by some other means and when the benefits of use strongly outweigh the negative consequences.

The Census Bureau continues to investigate the efficacy of web scraping. Current plans for web scraping focus on units, not people. Specifically, the Census Bureau is exploring the potential uses of internet data on residential permits and sales, state and local government data, and property taxes. However, some scraped unit data may include person information (e.g. property tax records listing the owners' names as an attribute of the property along with square footage, year built, etc.). This information could be used in processing to confirm appropriate linkages or to assess whether a unit is owner-occupant or rented. A panel of Division Chiefs familiar with ARTPD will consider the recommendation on justifying uses of PII from TPD.

5) Continue Gallup polling.

RESPONSE TO NAC ARTPD WG RECOMMENDATIONS

The Census Bureau has a solicitation out to conduct a Public Opinion Survey similar to the daily Gallup tracking poll that would allow us to continue monitoring public opinion surrounding the use of ARTPD. The Census Bureau expects to make an award this summer.

6) Rigorous ARTPD analyses are needed to consider the data source's reputation and data stewardship practices relative to those at Census, broadly defined costs, and benefits.

The Census Bureau uses memoranda of understanding (MOU), data use agreements, request letters, and legal contracts to acquire data. Regardless of format, the agreements are cleared and maintained by the Census Bureau Policy Office (MOUs and contracts are additionally cleared by Office of General Counsel at the Department of Commerce). While this clearance process does not explicitly consider the data source's reputation, we only enter into agreements with parties who are able and willing to agree to the terms and conditions required by the Census Bureau. Our understanding of the data source's data stewardship practices varies substantially across government and private entities.

To address the recommendation to analyze how data was collected by the source agency, the Census Bureau will ask each data source to describe where and how their information was gathered. We anticipate clearer responses from federal and state AR providers and vague responses from TPD providers.

The working group recommended that the costs of using ARTPD be analyzed, including the cost of acquiring data, cost of using ARTPD on data quality and coverage, cost in terms of reputation, and cost in terms of public trust. The Census Bureau is tracking and assessing costs to acquire and process ARTPD to improve census and survey operations. The cost in terms of data quality and coverage is being analyzed through simulations on 2010 Census data, in 2020 Census tests, and using the ACS. The costs in terms of reputation and public trust are being explored by the 2020 Census Privacy and Confidentiality Team and by the Communications Directorate-led group developing messaging strategies.

To address the recommendation to analyze benefits of ARTPD, the 2020 Census AR Modeling, Fitness for Use, Non-ID, and Contact Frame teams are documenting how and where ARTPD can improve operations, help with unit and item non-response, and reduce respondent burden.

7) Census Bureau should seek to acquire ARTPD for Hard-To-Reach/Hard-To-Count (HTC) populations that better covers those groups.

The Census Bureau is conducting market research to identify new AR and TPD data sources to improve person coverage. The Census Bureau will continue to engage stakeholder groups including Federal-State Cooperative for Population Estimates (FSCPE). Furthermore, the Census Bureau will continue to engage the NAC to discuss census policies, research and methodology, tests, operations, communications/messaging and other activities.

8) Census Bureau should look at ARTPD coverage of the HTC by area to determine if more targeted efforts are needed in certain parts of the country.

The Census Bureau regularly engages in community outreach programs to improve counts of HTC populations. Researchers have conducted geographic analyses of the coverage of AR (see the 2010 Census Match Study) and mail return rates (a measure of HTC populations) by county. These results were presented to the NAC working group at the March 2014 meeting.

9) Census Bureau should consider using statistical methods on ARTPD for data quality and coverage (e.g. imputation for missing data).

As with any data source, ARTPD may suffer from nonresponse. Nonresponse takes two forms: unit (e.g., omitted household) and item nonresponse (e.g., missing age). Depending on the type of nonresponse, the Census Bureau may employ weight adjustments, logical imputation, or statistical imputation to compensate. Whether the Census Bureau employs weighting or imputation techniques depends on the quality of the data, as well as its purpose. To address the Working Group's recommendation, the Census Bureau will continue to monitor and document ARTPD data quality to assess whether nonresponse adjustments would be appropriate.

10) Continued outreach and discussions with external data quality and privacy stakeholders.

The Census Bureau monitors media and publications on privacy and Big Data, and continues to engage stakeholders in the data quality and privacy fields. To improve data quality, the Census Bureau has ongoing efforts to seek out and obtain new data sources that may improve the coverage of ARTPD. Maintaining relationships with external data quality stakeholders is integral to these efforts to help the Census Bureau identify quality deficits. We will also maintain relationships with members of the NAC Privacy Subgroup, as well as contacts made through the member of the Privacy Subgroup, such as the privacy working group organized by The Information Technology and Innovation Foundation.

11) NAC should have a continuing advisory role as the Census Bureau continues to test and then integrate ARTPD for the 2020 Census.

The Census Bureau will provide updates on ARTPD plans for the 2020 Census at future NAC meetings.