Recommendations for OMB Statistical Policy on Tiered Access:
Developed at a COPAFS Sponsored Sloan Funded Workshop held
September 16, 2019

Recommendation 1.
We recommend that OMB encourage experimentation among different dimensions and options for access. In particular, we recommend that the FSRDC experiment on how it might adapt conditions for access, particularly considering the challenges of users who may not be associated with academic institutions. Promising examples of how this might be accomplished are being done, such as the pilot program on FSRDC remote access. Some key questions to consider in this research are: Why do you have to have a CB employee on-site? Why not cameras observed remotely? Why do researchers have to be Census Specially Sworn Employees? Wouldn’t some other contractual arrangement work?

Recommendation 2.
The Federal Statistical System requires statistical purpose for access to FSRDC. ‘Statistical purpose is defined by CIPSEA as the description, estimation or analysis of the characteristics of groups without identifying the individuals or organizations that comprise those groups. CIPSEA protected data is restricted from regulatory or enforcement use. Uses envisioned by the Foundations of Evidence-based Policy Act for program and policy evaluation, if defined appropriately, could meet the definition of statistical purpose. Any legislative barriers to such uses need to be identified and a plan developed to address them.

Recommendation 3.
In designating different tiers of access, a continuum for mitigating risk for each of the five safe constraints -- people, projects, setting, data, and output -- should be developed and applied.

Recommendation 4.
Research and development should continue into more promising solutions for access in lieu of public use microdata files, such as synthetic data, and the process for access. One suggestion is to use a researcher passport to streamline the process for applying for access to public-use data files of a type that were previously public.

Recommendation 5.
Research better risk quantification/estimation approaches. Westat has done risk estimation dozens of times successfully with at least 3 agencies using statistical modeling in the past 6 years and it produces some helpful information that can be used with other information to make recommendations for placing different products into different tiers.
Recommendation 6.

Conduct a holistic review of agency products to attempt to balance data utility with confidentiality and privacy constraints. This would include 1) identifying all the data sources (e.g., surveys), their data products, the main sources of risk, rules of disclosure (e.g., Rule of 3, etc), confidentiality edits (e.g., cell suppression rules), 2) for tables, determining if the table suppression system can be broken by the product itself or other data products, for a flexible table generator, determine if it is susceptible to differencing attacks, and for microdata estimate the risk. Recommendations can then be made for data architecture and risk mitigation.

Attendees at the Workshop

Julia Lane, NYU
Michael Davern, NORC
Peter Meyer, NORC
Tom Krenzke, Westat
Ken Hodges, Claritas
Margaret Levenstein, IPSR/UMI
Linda Jacobsen, PRB
Megan Gall, Leadership Conference on Civil Rights
Arloc Sherman, Center for Budget & Policy Priorities
Brian Harris-Kojetin, CNSTAT
Bill O’Hare, self
Nick Hart, Data Coalition
Brock Webb, Census Bureau/OMB Statistical Policy
Tim Mulhaney, NORC
Jennifer Madans, NCHS
Michael Hawes, Census Bureau
John Abowd, Census Bureau
Kristen Monaco, BLS
Heather Madray, FSRDC/Census Bureau
Tori Velkoff, Census Bureau
Shawn Bucholtz, HUD
Cynthia Clark, COPAFS
Corinna Turbes, COPAFS

Sponsored by

Alfred P. Sloan Foundation