

Statement of the American Economic Association's (AEA)

Committee on Economic Statistics <https://www.aeaweb.org/about-aea/committees/economic-statistics>
and

Committee on Government Relations <https://www.aeaweb.org/about-aea/committees/government-relations>

**on Desirable Characteristics of Repositories for Managing & Sharing Data from Federally Funded Research,
as invited in the Federal Register of January 17, 2020 (85 FR 3085)**

We commend the OSTP for having proposed a set of necessary characteristics of repositories for managing and sharing data from Federally-funded research, which would apply equally as well to data repositories for research funded by or originating from any source.

We also endorse the comments from the AEA Data Editor, Lars Vilhuber, submitted to you by separate package: <https://www.aeaweb.org/content/file?id=11689>, and wish to reiterate his opening comments on:

- The importance of sharing data (and computational instructions, “code”) for the purpose of transparency and reproducibility of science and the key role of data repositories in this endeavor; and
- The need for the scope of OSTP considerations to include research created by scientists in the direct employ of the federal government, data created for public and research use with federal funds as part of the business of the 13 federal principal statistical agencies, as well as any data created for research and evaluation under the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act).

Beyond this, we would appreciate some more explicit cross-walking to assure consistency between what OSTP is proposing and the work on data repositories and data access that are part of the Federal Data Strategy Action Plan, the 2019 guidelines issued by the Office of Management and Budget (OMB) to update the Information Quality Act <https://www.whitehouse.gov/wp-content/uploads/2019/04/M-19-15.pdf>, and the Evidence Act.

One feature of the Evidence Act on which OMB is working is the implementation of tiered data access to protected federal data. Tiered access can simultaneously widen access to data while limited the risk of nondisclosure by setting up a hierarchy of users. Low level users can access only a limited set of information, whereas the highest-level users can access the most sensitive data on the system. Incorporating this into the definition of OSTP’s *access* characteristic would demonstrate compatibility across Executive Branch efforts.

Another OMB and federal agency effort that may benefit from incorporation into OSTP’s proposals is the common application for data access being piloted by the Inter-university Consortium for Political and Social Research (ICPSR) under a contract with the U.S. Bureau of the Census. The pilot sets up a single portal and standard application process for requesting access to restricted data, across multiple data repositories or centers. This might be considered as an additional feature for OSTP’s *access* characteristic.

A final item that would benefit by clarification or coordination with OSTP’s proposal is the Evidence Act’s establishment of a Data for Evidence Advisory Committee charged with setting up a National Secure Data Service facilitating access to all government data. What might be the similarities and differences between a set of data repositories as prescribed by OSTP and a National Secure Data Service?

Clearing up confusion among similar national efforts within the Executive Branch would be beneficial to organizations that, like the AEA, wish to support and facilitate federal efforts that improve data storage, curation, and access for research purposes.

Thank you for the opportunity to comment.

John C. Haltiwanger
Chair, Committee on Economic Statistics

Kenneth Troske
Chair, Committee on Government Relations