Congressional Research Service
Informing the legislative debate since 1914

IN FOCUS

November 20, 2026

Preserving Trust: Government Data, Disclosure Avoidance, and Privacy Enhancing Techniques (PETs)

The federal government collects and maintains data on individuals to use in conducting authorized governmental functions. It may use administrative data to facilitate delivery of programs and statistical data to better inform the nation's understanding of groups within society. For decades, federal agencies have recognized the challenge between maintaining sufficiently detailed data on individuals to efficiently perform government functions and protecting individuals' privacy by allowing them to understand and consent to how their data are being used. While sharing data across agencies can streamline governmental processes, doing so without sufficient safeguards can also increase the likelihood that data are used for unintended purposes.

In 1973, Congress enacted the Privacy Act of 1974 (5 U.S.C. §552a) to regulate the federal government's use of individually identifiable information. The Privacy Act generally requires agencies to adhere to certain notice and consent requirements when working with individually identifiable information. In 2002, Congress enacted the Confidential Information Protection and Statistical Efficiency Act (CIPSEA; 44 U.S.C. §§3561-3583) for similar reasons to require that agencies pledge to keep statistical data confidential to preserve respondent trust and candidness.

Since enactment of these laws, questions have emerged concerning the ability of agencies to protect privacy when working with administrative and statistical data, especially as computing technologies continue to advance and raise new questions about data protection that were not apparent in the pre-internet era. Questions may include how to manage new privacy risks associated with an enhanced ability to access, combine, and use confidential and public data. In response, agencies have explored and implemented new disclosure avoidance methods and privacy enhancing techniques (PETs, also referred to as privacy enhancing technologies).

Trust: Weighing Privacy and Accuracy

The Office of Management and Budget's (OMB's) Statistical Policy Directive No. 1 explains that "any loss of trust in the relevance, accuracy, objectivity, or integrity of the Federal statistical system and its products can foster uncertainty about the validity of measures our Nation uses to monitor and assess performance, progress, and needs." If public trust is lost, support for statistical surveys may decline. This decline in trust may result in lower response rates and inaccurate or truncated responses from individuals surveyed. If public trust is lost, the data resulting from any survey may become a less reliable reflection of the population surveyed.

At the same time, the values of privacy and accuracy of data often share an inverse relationship. Providing the most amount of data is often the most accurate and useful method for research and program administration, but as the level of detail increases, it also creates more risks that such data can be tied to specific persons. The respective weighting of privacy and accuracy may be thought of as a range. For example, in uses where the information is more sensitive, such as researching health outcomes in rural areas, users can prioritize certain data fields and remove others to increase privacy while sacrificing detail. In a different context where the data are less sensitive, such as in research concerning incentives to promote hiring by small businesses, users can include more fields to preserve the accuracy and usefulness of the dataset for future purposes at the expense of a degree of privacy.

The Mosaic Effect and New Privacy Risks

Privacy may be compromised not only through direct disclosures of identifying information but also indirectly through the *mosaic effect*, first described by OMB in 2013, where seemingly secured, deidentified data can be combined with other datasets to reidentify individuals. As OMB explained in 2025, "The mosaic effect occurs when the information in an individual data asset, in isolation, may not pose a risk of, for example, identifying an individual, but when combined with other available information, could pose such risk" (OMB Memorandum M-25-05). For example, communities with small populations—such as individuals within rural areas, certain racial categories, or outlying income groups—may be more readily identifiable from data where insufficient disclosure avoidance methods have been applied.

Disclosure Avoidance Methods

To preserve public trust and comply with statutory privacy requirements, agencies may choose between different methods to obscure data and preserve privacy while also retaining the usability of the data for different administrative and statistical use cases. These *disclosure avoidance* (or limitation) methods are intended as one way to preserve trust by hindering anyone from identifying individual respondents from available information.

Agencies have historically employed various disclosure avoidance methods to protect and obscure individually identifying information (e.g., name, education, and medical history) when disseminating data products. Over time, these methods have expanded from traditional approaches (such as *coarsening*, which refers to the rounding or aggregation of data to reduce detail, and *suppression*, which refers to the redaction of data) to include PETs. According to the Census Bureau, these traditional methods can introduce accuracy

concerns by reducing the amount of data that can be supplied and distorting population counts.

Unlike traditional approaches, PETs are often characterized by their use of advanced computing and statistical techniques. PETs can be grouped by their different characteristics, such as whether the PET obfuscates information at the data level (such as differential privacy or synthetic data), encrypts data before sharing limited information to vetted users (as in secure multiparty computation or homomorphic encryption), or provides an isolated computing environment for sensitive data parallel to an agency's main operating system (such as in trusted execution environments). Just as traditional approaches entail certain trade-offs, PETs can vary in terms of their value for preserving privacy and usability, their cost and scalability for certain applications, and whether they can be applied only to small or simple datasets. They can also vary in terms of whether and how they manipulate the originally collected data, each having different impacts on the potential accuracy of the data or the resulting computations.

OMB Memorandum M-25-05 requires agencies to assess reidentification risk from other publicly available data and use disclosure avoidance methods like suppression and PETs to manage the risks. However, when agencies use techniques like differential privacy or noise injection that knowingly result in inaccurate data, OMB requires agencies to provide certain accompanying documentation regarding how the technique was applied and its limitations, describing how users can access the unaltered data, and providing a clear indication that the data have been altered only to meet privacy requirements and not for other purposes.

Case Study: The U.S. Census Bureau's Use of Disclosure Avoidance Methods Over Time

The U.S. Census Bureau provides an illustrative example of how agencies' use of disclosure avoidance methods can change over time. In the 1970 and 1980 censuses, the Census Bureau primarily relied on suppressing data tables from publication. Data users expressed dissatisfaction with the usability and replicability of what was provided, and the Census Bureau moved to different disclosure avoidance methods because of the low utility of the suppressed data tables. In the 1990, 2000, and 2010 censuses, the Census Bureau expanded use of *swapping*, where sensitive characteristics were exchanged at the household level to disguise the true respondent answers.

As computing technologies continued to advance, studies from the early 2000s found that individuals could be reidentified by linking shared attributes and that 87% of the population in 2000 was likely to be uniquely identifiable. These attributes include zip code, birth date, and sex, which were available from voting history or medical information. Notwithstanding some concerns from scholars, the Census Bureau moved away from previous disclosure avoidance methods and adopted *differential privacy* for the 2020 census. A corresponding Census Bureau report explained that *differential privacy* is a mathematical framework for quantifying the level of disclosure risk. According to the Census Bureau, through this framework, agencies can track potential privacy loss as they publish more data, reducing the risk of identifying an individual but allowing the manipulated data to perform similarly to the originally collected data, if used appropriately.

As part of the Census Bureau's periodic assessments of previous use of disclosure avoidance methods, in 2025, the Census Bureau published the results of a simulated reidentification attack on the 2010 data, finding that an attacker could reidentify all records in 70% of all census blocks (97 million people) using only published data. The same paper found that the methods used for the 2020 data, which included PETs, more robustly protected against such reconstruction attacks. At a November 2025 hearing, held by the Subcommittee on the Constitution and Limited Government of the House Committee on the Judiciary, some witnesses raised concerns about the impact of differential privacy on accuracy; others noted that previous swapping methods may not secure privacy as computing power and digital information availability increase.

On June 4, 2026, the Department of Commerce limited the Census Bureau's and Bureau of Economic Analysis's use of disclosure avoidance methods to coarsening and suppression. The department disallowed techniques using noise infusion, such as differential privacy, to provide "the public with accurate statistical information." Several professional statistical associations issued a joint statement on the June order, expressing concern that the limitation will lead to less privacy for personal information collected by these agencies, that the data will be less usable, or both. Findings from other studies seem to call into question whether differential privacy and noise injection have any impact on the accuracy of statistical products.

Assessing Privacy and Accuracy Trade-Offs

In response, Congress may consider whether statutes like the Privacy Act or CIPSEA are sufficiently specific to weigh privacy interests against the accuracy of information or whether additional research is needed to determine an agency-wide standard. Congress may also examine the implementation of Memorandum M-25-05 guidance with respect to the sufficiency of agency explanations for using disclosure avoidance methods and PETs.

When evaluating an agency's chosen disclosure avoidance method, such as the Department of Commerce's recent decision to return to traditional coarsening and suppression disclosure avoidance methods, Congress could inquire about how the decision was reached and what information on the use of alternative methods, such as PETs, was collected and assessed before making the decision. Congress could also require the Department of Commerce, OMB, or National Institute of Standards and Technology to perform additional research on the comparative advantages and disadvantages of disclosure avoidance methods with respect to both accuracy and privacy.

Meghan M. Stuessy, Analyst in Government Organization and Management

IF13358

Disclaimer

This document was prepared by the Congressional Research Service (CRS). CRS serves as nonpartisan shared staff to congressional committees and Members of Congress. It operates solely at the behest of and under the direction of Congress. Information in a CRS Report should not be relied upon for purposes other than public understanding of information that has been provided by CRS to Members of Congress in connection with CRS's institutional role. CRS Reports, as a work of the United States Government, are not subject to copyright protection in the United States. Any CRS Report may be reproduced and distributed in its entirety without permission from CRS. However, as a CRS Report may include copyrighted images or material from a third party, you may need to obtain the permission of the copyright holder if you wish to copy or otherwise use copyrighted material.