epic.org

Electronic Privacy Information Center
1519 New Hampshire Avenue NW
Washington, DC 20036, USA

+1 202 483 1140
+1 202 483 1248
@EPICPrivacy
https://epic.org

COMMENTS OF THE ELECTRONIC PRIVACY INFORMATION CENTER

to the

NATIONAL INSTITUTE of STANDARDS AND TECHNOLOGY

Request for Comment on De-Identifying Government Data Sets Paper (3rd Draft)

January 13, 2023

By notice published on November 15, 2022, the National Institute of Standards and Technology (NIST) have requested comment on its De-Identifying Government Sets paper (3rd draft) which provides guidance on de-identification techniques, process, and implementation to governance agencies.[1] Notably, this third draft paper reflects advancement since the previous draft six years ago, as "there has been significant developments in privacy technology, specifically in the theory and practice of differential privacy."[2]

The Electronic Privacy Information Center (EPIC) submits these brief comments to commend NIST's revisions and to recommend further changes. EPIC is a public interest research center in Washington, D.C., established in 1994 to focus public attention on emerging privacy and related human rights issues and to protect privacy, the First Amendment, and constitutional values.[3] EPIC has a longstanding interest in federal efforts to develop privacy-enhancing technologies and regularly comments on federal planning efforts at the intersection of technology and privacy.[4] EPIC

[1] NIST, De-Identifying Government Data Sets (3rd Draft) Request for Comment (Nov. 15, 2022), https://csrc.nist.gov/publications/detail/sp/800-188/draft.
[2] *Id*.
[3] EPIC, *About Us* (2022), https://epic.org/about/.
[4] *See, e.g.*, Comments of EPIC, Request for Information on Federal Video and Image Analytics Research and Development Action Plan, 87 Fed. Reg. 42,212 (Sept. 2, 2022), https://epic.org/documents/epic-comments-in-re-federal-video-and-image-analytics-research-development-action-plan/; Comments of EPIC, Public and

has repeatedly intervened in support of robust privacy safeguards for personal information contained in government records,[5] including the Census Bureau's adoption of differential privacy for its disclosure avoidance system.[6] In recent comments to the White House Office of Science and Technology Policy, EPIC urged federal agencies to prioritize the adoption of differential privacy and increase funding across the board for privacy-enhancing techniques.[7]

The risk of reidentification from nominally deidentified data sets is profound and growing, posing serious risks to the public. Formal privacy techniques like differential privacy should be implemented to counter the harms that can result from the disclosure of personal information through government data sets. As NIST rightly notes, these harms can include embarrassment or injury to an individual, disclosure of sensitive information, inferential disclosure concerning an entire class of individuals,[8] and increased commercial surveillance harms facilitated by commercially available deidentified data, like high-resolution geolocation information.[9]

Private Sector Uses of Biometric Technologies, Office of Sci. & Tech. Policy (Jan. 15, 2022), https://epic.org/documents/epic-comments-to-ostp-on-public-and-private-sector- uses-of-biometric-technologies/; Comments of EPIC, Artificial Intelligence Risk Management Framework, Nat'l Instit. of Standards & Tech. (Aug. 18, 2021), https://epic.org/documents/regarding-the-artificial- intelligence-risk-management-framework/; Comments of EPIC, Request for Information (RFI) on an Implementation Plan for a National Artificial Intelligence Research Resource, Office of Sci. & Tech. Policy & Nat'l Sci. Found. (Oct. 1, 2021), https://epic.org/wp-content/uploads/2021/10/EPIC-Comment-NAIRR- Oct2021.pdf; Comments of EPIC, Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning, Comptroller of the Currency et al., (July 1, 2021), https://archive.epic.org/apa/comments/EPIC-Financial-Agencies-AI-July2021.pdf; Comments of EPIC, Solicitation of Written Comments by the National Security Commission on Artificial Intelligence, 85 Fed. Reg. 32,055 (Sep. 30, 2020), https://epic.org/wp-content/uploads/apa/comments/EPIC-comments-to-NSCAI-093020.pdf; Comments of EPIC, Request for Comments on a Draft Memorandum to the Heads of Executive Departments and Agencies, "Guidance for Regulation of Artificial Intelligence Applications," (Mar. 13, 2020), https://epic.org/apa/comments/EPIC-OMB-AI-MAR2020.pdf.

[5] EPIC, *Data Protection: Government Records and Privacy* (2022), https://epic.org/issues/data-protection/government-records-privacy/.

[6] EPIC, *Census Privacy* (2022), https://epic.org/issues/democracy-free-speech/census-privacy/.

[7] https://epic.org/documents/epic-comments-to-ostp-on-advancing-differential-privacy/

[8] Simson Garfinkel et al., *De-Identifying Government Data Sets: Third Public Draft*, NIST 13 (November 2022), https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-188.3pd.pdf.

[9] *Id*. at 1.

EPIC commends NIST for providing detailed and thoughtful guidance to a federal agencies about deidentification processes, formal privacy models, and implementation strategies for Disclosure Review Boards. The draft paper identifies key factors for a Disclosure Review Board to evaluate risk and the probability of reidentification, and we anticipate that it will be a valuable resource for mitigating the harms that can flow from government data sets in an era of increasingly sophisticated reidentification techniques.

However, we urge NIST to take its guidance a step further by (1) more forcefully encouraging the adoption of differential privacy over prescriptive deidentification standards whenever possible, and (2) identifying specific circumstances where differential privacy should be used or strongly favored. As NIST correctly notes, assurances concerning the risk of reidentification frequently cannot be made without the use of formal privacy techniques—principally, differential privacy. Although the draft paper provides agencies with a menu of deidentification tools and techniques, NIST should use its unique standing and credibility to urge agencies to adopt differential privacy as a mathematically rigorous technique for measuring and controlling the risk of reidentification.

EPIC thanks NIST for its attention to these issues and for taking the time to consider EPIC's recommendations. If you have any questions, please feel to contact EPIC Fellow Suzanne Bernstein at bernstein@epic.org.

Sincerely,

/s/ John Davisson
John Davisson
EPIC Director of Litigation &
Senior Counsel

/s/ Suzanne Bernstein
Suzanne Bernstein
EPIC Fellow