



Privacidade, confidencialidade e confiança: lições do debate sobre microdados censitários para o futuro das estatísticas públicas

Eduardo L. G. Rios-Neto*

Introdução

O microdado de uso público constituiu infraestrutura essencial para a pesquisa demográfica durante décadas. A amostra de 1 em 1.000 registros do Censo norte-americano de 1960 permitiu aos pesquisadores formular hipóteses após a coleta dos dados, definir subpopulações conforme questões de pesquisa e construir comparações retrospectivamente, sem se limitar às tabulações previamente definidas pela agência. O microdado, portanto, ia além da estatística, constituindo também uma infraestrutura para a ciência social aplicada.

Esse modelo entrou em crise com o aumento da capacidade computacional, a proliferação de fontes externas e a combinação de grandes volumes de dados publicados. O caso norte-americano é particularmente relevante porque a controvérsia foi documentada, submetida à evidência empírica, discutida em tribunais e confrontada pela comunidade científica. Essa experiência oferece elementos importantes para a reflexão sobre decisões semelhantes em outros sistemas estatísticos.

Três questões devem ser tratadas separadamente no âmbito da controvérsia. A primeira é de natureza metodológica: como proteger indivíduos em um ambiente no qual a reidentificação pode ser concebida como um problema estrutural? A segunda é de natureza científica: quanta informação pode ser removida, agregada ou perturbada sem que o microdado deixe de cumprir sua função precípua no estudo científico? A terceira questão é institucional: quem decide sobre o equilíbrio entre proteção e utilidade, segundo quais critérios e com que participação dos usuários?

* Universidade Federal de Minas Gerais (UFMG), Belo Horizonte-MG, Brasil (riosnetoedu@gmail.com; <https://orcid.org/0000-0002-3161-5791>).

Da proteção tradicional à privacidade diferencial

A experiência norte-americana mostra que proteção e utilidade foram acomodadas sucessivamente. Ruggles e Magnuson (2023), em uma reconstrução histórica do período de 1790 a 2020, distinguem os conceitos de privacidade da confidencialidade. Privacidade diz respeito ao controle do indivíduo sobre a circulação de suas informações, enquanto confidencialidade refere-se à promessa estatal de não revelar dados pessoais. Essa distinção é importante porque boa parte da política censitária respondeu historicamente à preocupação com a intrusão do Estado, reforçando a confidencialidade, sem necessariamente resolver a questão original do uso das informações.

A partir de 1930, o Census Bureau deixou de publicar dados de áreas muito pequenas e, a partir de 1970, passou a suprimir tabelas inteiras. A insatisfação dos usuários com os custos analíticos dessas medidas contribuiu, nos anos 1990, para a adoção da troca de registros (*data swapping* ou *data switching*), utilizada até 2010. As mudanças de método foram motivadas não apenas pela segurança, mas também pelos possíveis prejuízos à análise científica.

A troca de registros visa a gerar incerteza quanto à associação entre a pessoa, o atributo e a localização. Para isso, valores de determinados campos são permutados entre registros semelhantes, de modo que um registro divulgado não possa ser associado com segurança à pessoa ou à área de origem. No desenho descrito para o Censo norte-americano, o pareamento baseava-se em características como tamanho do domicílio, raça, origem hispânica e número de adultos, preservando as estatísticas construídas a partir dessas variáveis e as contagens populacionais nos níveis geográficos. Sua vantagem é a seletividade, uma vez que a proteção pode concentrar-se nos registros considerados mais vulneráveis. Entretanto, a taxa de troca não era divulgada, o que representava uma fragilidade estrutural. Se a eficácia depende de parâmetros secretos, o usuário externo não consegue auditar o método nem corrigir suas estimativas para mitigar um possível viés. Assim, embora a proteção pode ser efetiva, sua magnitude e seu custo estatístico permanecem difíceis de mensurar.

A posição do Census Bureau mudou substancialmente com a teoria da reconstrução. O resultado apresentado por Dinur e Nissim (2003) mostra que a publicação de um conjunto suficientemente rico de estatísticas sobre o mesmo universo pode, em princípio, permitir a reconstrução dos registros subjacentes. A cada nova tabulação acrescentam-se restrições ao sistema e, a partir de determinado ponto, um mecanismo que parecia seguro quando considerado isoladamente pode tornar-se vulnerável quando as divulgações são analisadas em conjunto. Abowd e Schmutte (2019) transformaram essa proposição em um argumento empírico. Com base em mais de 7,7 bilhões de sumários publicados no Censo de 2010, o Bureau reconstruiu atributos como quarteirão, idade, sexo, raça e origem hispânica e tentou vinculá-los a bases de dados comerciais. Para a agência, a vulnerabilidade deveria ser avaliada antes de qualquer incidente, e não apenas a partir da existência de

reidentificações documentadas. Nessa perspectiva, a privacidade diferencial oferece uma garantia matemática *ex ante*, com parâmetros publicáveis, resistência a determinadas formas de pós-processamento e uma estrutura para contabilizar a perda de privacidade ao longo de publicações sucessivas.

Abowd e Schmutte (2019) tratam o parâmetro épsilon como parte de uma escolha social: não existe uma fronteira tecnicamente única entre privacidade e acurácia. Escolher um nível de proteção significa decidir quanta precisão estatística a sociedade está disposta a sacrificar para reduzir determinado risco. O parâmetro técnico incorpora preferências sobre o uso futuro da informação, inclusive dimensões econômicas e políticas.

A crítica dos usuários: risco verificável e dano à utilidade

Ruggles (2024) propõe um padrão de prova distinto. Primeiro, questiona o experimento de reconstrução do Census Bureau, argumentando que as tabelas do Censo de 2010 já haviam passado por uma troca de registros. Assim, a reconstrução não teria recuperado simplesmente os microdados originais. Aliás, o próprio Bureau reconheceu que mais da metade das reidentificações putativas estava incorreta. A crítica central é que, se um atacante externo não consegue distinguir os casos corretos dos incorretos, a taxa de reidentificação putativa não constitui, por si só, uma medida de ameaça explorável.

Há, portanto, uma divergência entre a agência e os usuários. Para a agência, a existência de uma vulnerabilidade estrutural é suficiente para justificar a adoção de mecanismos formais de proteção. Para os usuários, a política deve demonstrar a ocorrência de dano verificável e considerar a probabilidade de exploração. O desacordo, portanto, não é apenas quantitativo: envolve o que conta como evidência suficiente para alterar uma infraestrutura científica consolidada.

O segundo ponto da crítica diz respeito às consequências sobre a utilidade estatística. Os arquivos de demonstração da nova abordagem, divulgados entre 2019 e 2021, foram avaliados externamente e considerados inadequados para diversos usos críticos. O pós-processamento destinado a impor a não negatividade e a consistência hierárquica teria produzido padrões sistemáticos de erro, afetando populações urbanas, suburbanas e rurais, minorias em diferentes contextos de concentração, estimativas de migração, taxas de mortalidade por Covid-19 e medidas de segregação residencial. Não se trata apenas de aumentar a variância: determinadas populações e áreas geográficas podem ser afetadas sistematicamente.

A evidência independente não converge para uma única conclusão, o que recomenda avaliação específica de cada produto estatístico. Uma primeira linha de avaliação do algoritmo finalizado mostrou que as discrepâncias em relação ao método anterior crescem de forma sistemática ao longo do contínuo urbano–rural e afetam de maneira desigual os grupos não brancos, comprometendo não apenas as contagens, mas também as taxas de crescimento intercensitárias. Uma segunda avaliação, apoiada no arquivo de medidas

ruidosas divulgado em 2023, chegou a um resultado parcialmente contrário: o ruído bruto é excessivo para uso direto, sobretudo para hispânicos e multirraciais, mas o pós-processamento do TopDown restitui acurácia comparável à da troca de registros, com erros, em geral, não superiores aos de outras fontes de erro censitário, exceto em áreas de população pequena. Uma terceira linha de investigação sustenta que a troca dirigida protege melhor justamente os respondentes mais vulneráveis e que o risco de reidentificação atribuído às divulgações anteriores foi superestimado devido a um procedimento inadequado de pareamento. Uma quarta avaliação, enfim, demonstra que as garantias efetivamente entregues pelo Censo de 2020 são mais fortes do que as publicadas, de modo que a variância do ruído poderia ter sido reduzida entre 15,1% e 24,8% sem perda relevante de proteção. Essas avaliações não convergem quanto às conclusões, mas compartilham a condição de possibilidade: quando arquivos, parâmetros e métricas são disponibilizados, a contabilidade entre risco e utilidade pode ser refeita por terceiros (Mueller; Santos-Lozada, 2022; Kenny *et al.*, 2024; Muralidhar *et al.*, 2025; Su; Su; Wang, 2025).

Ruggles (2024) propõe, nesse contexto, o uso de instrumentos seletivos. A troca dirigida poderia concentrar-se em registros efetivamente vulneráveis, enquanto a infusão de ruído afetaria inclusive células em que não há combinações únicas a proteger. A experiência do Office for National Statistics, no Censo de 2021 da Inglaterra e do País de Gales, que combinou troca dirigida e perturbação por chave de célula, sugere uma conclusão menos binária: diferentes métodos de proteção podem ser complementares e devem ser avaliados de acordo com as características e finalidades de cada produto estatístico, em vez de serem impostos como solução institucional única (ONS, 2023).

O problema dos microdados sintéticos

A controvérsia em torno do ACS é ainda mais profunda quando considerada do ponto de vista da pesquisa científica. O arquivo público da pesquisa descreve, anualmente, milhões de pessoas e domicílios, mantém uma série histórica desde 1962 e sustenta uma ampla comunidade de pesquisadores. Sua função não se limita a permitir a reprodução de análises já realizadas; consiste também em preservar a possibilidade de descobrir relações que não haviam sido previamente formuladas. É nesse ponto que se situa uma crítica de fundo aos microdados sintéticos. Um arquivo sintético reproduz as relações entre variáveis que o modelo gerador conseguiu representar. Embora possa ser adequado para replicar análises previamente concebidas, ele é potencialmente inadequado para explorar combinações inesperadas, subpopulações pequenas ou especificações não previstas quando o modelo foi ajustado (Ruggles, 2025).

A participação dos usuários tornou-se parte da governança desse processo. Em 2019, milhares de pesquisadores de centenas de instituições ofereceram ao Bureau a replicação sistemática de estudos revisados por pares sobre os dados protegidos. A proposta buscava transformar a discussão sobre a utilidade dos dados em uma avaliação empírica. A

ausência de uma resposta direta contribuiu para que uma comunidade de pesquisadores antes dispersa se constituísse como um ator relevante no debate sobre a política de divulgação dos dados.

A utilidade de um microdado não se mede apenas pela capacidade de reproduzir tabelas já existentes. Seu valor científico reside também na possibilidade de formular e testar perguntas ainda não concebidas. Quanto mais a proteção incorpora um modelo sobre quais relações são importantes, maior é o risco de transformar a base de dados em um mecanismo de confirmação de hipóteses já estabelecidas.

A discussão das seções anteriores resume-se no Quadro 1.

QUADRO 1
Estrutura comparada das duas posições

Dimensão	Abowd / Census Bureau	Ruggles / comunidade de usuários
Fundamento	Teorema da reconstrução: tabulação abundante é matematicamente invertível	Dois séculos de prática censitária sem prejuízo verificado
Padrão de prova	Garantia matemática <i>ex ante</i> , independente de incidente observado	Prejuízo real, comprovável e verificável <i>ex post</i>
Unidade de risco	O sistema de divulgação como um todo	O indivíduo efetivamente exposto
Solução	Privacidade diferencial formal, uniforme sobre a hierarquia geográfica	Métodos tradicionais aprimorados, dirigidos ao risco caso a caso
Posição sobre o ACS	Migração futura para microdados totalmente sintéticos	Oposição frontal: perda irreparável de descoberta não antecipada

Fonte: Abowd (2021), Abowd *et al.* (2022), Abowd e Schmutte (2019), Ruggles (2024) e Ruggles (2025).

Nota: O mesmo experimento é interpretado de modo oposto conforme o padrão de prova.

A inversão institucional de 2025-2026 nos Estados Unidos

Os acontecimentos de 2025-2026 alteraram o problema. A controvérsia deixou de ser apenas metodológica e passou a envolver diretamente a agência, seu órgão supervisor e o universo estatístico produzido.

Em junho de 2026, o Departamento de Comércio editou uma ordem administrativa sobre controle de divulgação que proibiu o uso de infusão de ruído em produtos estatísticos. A medida afetou imediatamente a divulgação de estimativas do ACS e, segundo organizações usuárias, tende a reduzir estatísticas disponíveis para pequenas áreas e populações de menor dimensão. O episódio chama a atenção: a mesma linguagem de inteligibilidade e confiança pública que aparecia na crítica à privacidade diferencial passou a justificar a restrição de uma técnica de proteção. Sem investimento suficiente em modalidades de acesso controlado, a eliminação do ruído pode deslocar o problema para a supressão: as células publicadas tornam-se mais fiéis, mas parte da informação deixa de ser publicada. A questão, portanto, não é apenas qual técnica oferece maior proteção, mas também quem decide, segundo quais procedimentos e com que evidência.

Em 10 de setembro de 2026, ainda sob consulta pública com prazo até 13 de outubro, o Census Bureau propôs uma regulamentação que redefine a residência habitual: apenas cidadãos e residentes permanentes legais seriam contados para o *apportionment*, e o *status* migratório passaria a ser apurado por meio de registros administrativos — fiscais, migratórios, postais e comerciais — e, eventualmente, por meio de um quesito direto no questionário. Isso rompe com mais de dois séculos de prática: há quesito de raça em todos os censos desde 1790, e o estrangeiro residente sempre integrou a base de *apportionment* (U.S. Census Bureau, 2026; U.S. Dep. of Commerce, 2026).

A análise histórica referenciada no debate anterior mostra que a desconfiança dos respondentes está mais ligada ao receio de uso indevido da informação pelas autoridades do que ao risco técnico de reidentificação. Essa consideração ganha especial relevância quando a própria enumeração passa a depender de registros administrativos relacionados ao *status* migratório.

Esse episódio indica uma conclusão metodológica: padrões de transparência considerados indispensáveis para dados estatísticos ordinários não podem ser abandonados quando o resultado é politicamente sensível. A autoridade de uma agência estatística também decorre da regularidade dos procedimentos. Quando autoria, método, revisão e mensuração do erro desaparecem, o problema deixa de ser apenas a qualidade de uma estimativa e passa a afetar a credibilidade de todo o sistema. Cabe registrar, porém, uma diferença de procedimento: diferentemente da ordem administrativa de junho, a proposta de setembro está submetida a consulta pública, com prazo e registro formal. O procedimento existe; sua eficácia dependerá de como a agência responderá ao que o registro produzir.

O caso brasileiro: Censo Demográfico 2022

A experiência brasileira oferece um contraste instrutivo. As *Notas Metodológicas 03/2026* do IBGE formalizam uma mudança no regime de disseminação dos microdados da amostra. Nos censos de 2000 e 2010, havia um produto único, com *download* gratuito e sem identificação do usuário, que tinha a “Área de Ponderação” como menor nível geográfico, além de idade em anos simples e pesos consistentes com as tabulações oficiais. Em 2022, o IBGE adotou uma arquitetura em três modalidades.

No acesso público, a unidade da federação passa a ser o menor nível geográfico; a idade é recodificada em grupos quinquenais, com topo em 80 anos; uma subamostra de 50% dos domicílios é utilizada em setores de fração amostral integral, com supressão de cerca de 116 mil domicílios; aplicam-se supressões locais e globais; e são retiradas variáveis sem identificadores com muitas categorias. O acesso controlado restitui a área de ponderação e parte das variáveis mediante autenticação, termo de compromisso, declaração de vínculo institucional e rastreabilidade individual. Já o acesso restrito, realizado na Sala de Acesso a Dados Restritos, permite maior detalhamento territorial e o acesso aos pesos das tabulações oficiais, mediante projeto submetido ao comitê responsável.

Essa arquitetura é prudente e inovadora institucionalmente. O IBGE não adotou nem a infusão de ruído nem os microdados sintéticos como solução universal; deslocou parte relevante da proteção para a arquitetura de acesso. Essa lógica de camadas converge com a recomendação presente no debate internacional: diferentes riscos e produtos podem exigir níveis distintos de proteção. A experiência norte-americana sugere três alertas. Primeiro, supressão e agregação também têm efeitos distributivos. Ao eliminar detalhes, essas técnicas tendem a afetar as caudas da distribuição, nas quais se encontram justamente pequenas áreas e populações de menor dimensão. A diferença em relação ao ruído é que o erro da supressão é determinístico e, por isso, pode ser mais transparente, mas não necessariamente menos relevante. Segundo, há uma decisão de política estatística apresentada como parâmetro técnico: escolher o que será preservado significa escolher quais análises terão maior estabilidade. Toda arquitetura de proteção precisa explicitar quais propriedades estatísticas permanecerão verdadeiras. Terceiro, permanece uma lacuna de transparência. Conforme o documento analisado, as notas não publicam o limite de risco adotado nem a métrica de risco individual correspondente. Tampouco há, até o material considerado, arquivos de demonstração ou de avaliação independente sistemática do impacto sobre estudos consolidados. Além disso, restrições à redistribuição podem criar tensões com as práticas de reprodutibilidade científica.

Há que se ter cautela na comparação entre o Brasil e os Estados Unidos. No Brasil, a mudança é apresentada como uma decisão técnica do próprio IBGE, documentada e fundamentada na legislação de proteção de dados. Nos Estados Unidos, parte das mudanças recentes veio de fora da agência, acompanhada de alterações no ambiente institucional. Não são situações equivalentes, distinção essencial para uma análise justa.

Discussão: três decisões que não devem ser confundidas

A controvérsia permite um quadro decisório mais amplo. A primeira decisão diz respeito ao padrão de prova de risco. Uma agência pode exigir uma demonstração estrutural *ex ante*, como no argumento da reconstrução, ou priorizar evidências históricas e empíricas de dano, como na crítica dos usuários. As duas abordagens respondem a perguntas distintas: a primeira procura determinar se uma vulnerabilidade é possível em princípio; a segunda, se ela é explorável e relevante nas condições reais do produto. Uma política de proteção madura precisa explicitar o padrão de prova aplicado, em vez de alterná-lo conforme a conveniência do argumento. A segunda decisão refere-se à unidade de risco. O sistema de divulgação pode ser considerado como um todo, acumulando todas as estatísticas publicadas, ou a análise pode se concentrar no indivíduo efetivamente exposto. A escolha altera a avaliação. Um mecanismo que parece necessário para proteger o sistema inteiro pode impor custos desproporcionais a populações que não apresentam risco relevante no produto específico. A terceira decisão diz respeito ao custo científico aceitável. Não há utilidade nem privacidade abstratas. Há produtos estatísticos destinados a finalidades

específicas, com usuários distintos e riscos diversos. A proteção deve ser proporcional ao uso previsto, não necessariamente idêntica em toda a hierarquia de produtos.

Essa abordagem favorece uma arquitetura híbrida. Perturbação, troca, supressão, agregação e acesso controlado não precisam ser mutuamente exclusivos. Podem combinar-se conforme risco, produto e finalidade, desde que seus efeitos sobre a utilidade sejam mensuráveis. É nesse ponto que o teste de intruso se torna indispensável. Nesse exercício, analistas tentam reidentificar registros com base no produto divulgado e em fontes externas. O teste liga o risco teórico ao comportamento observável. Não substitui garantias matemáticas, mas acrescenta evidência auditável. O teste deve ser periódico, sobretudo quando a agência altera sua política de disseminação, com publicação de critérios, limitações e resultados. Arquivos de demonstração também devem preceder decisões definitivas. A participação não deve ser simbólica. A comunidade usuária conhece usos que a agência não consegue antecipar integralmente. A participação deve produzir registros verificáveis: estudos replicados, problemas identificados, respostas da agência.

Conclusão: confiança como infraestrutura estatística

O ponto de vista termina com duas incertezas: técnica e institucional. Tecnicamente, ainda há dificuldade para traduzir parâmetros de privacidade diferencial em risco absoluto de divulgação, para aplicar garantias formais a microdados de propósito geral sem degradação excessiva, para preservar garantias quando invariantes derivadas dos dados confidenciais são impostas, para quantificar o risco acumulado da troca de registros, para validar dados sintéticos em relações não antecipadas pelo modelo e para medir os efeitos distributivos da supressão. Nenhuma dessas questões deve ser minimizada. Institucionalmente, a estatística oficial constitui um bem cuja produção depende da confiança. O usuário não consegue verificar diretamente a maior parte do processo de produção e precisa confiar na instituição que mede, documenta e divulga os dados. Autoria, revisão, documentação metodológica, mensuração do erro e consistência histórica são instrumentos que constroem a confiança institucional. A confiança do respondente e a do usuário também estão relacionadas. Se o público suspeita do uso indevido da informação por autoridades, a disposição para responder pode diminuir, a qualidade do dado pode se deteriorar e a reputação da agência pode ser afetada. A proteção formal, sem considerar esse mecanismo, pode consumir utilidade sem gerar maior confiança social.

Para o sistema estatístico brasileiro, isso significa que a arquitetura em camadas adotada para o Censo 2022 pode ser fortalecida por cinco compromissos: publicar o padrão de risco e sua métrica; disponibilizar arquivos de demonstração antes das decisões finais; convidar usuários a replicar estudos consolidados; registrar formalmente o padrão de prova aplicado em cada liberação; e reavaliar periodicamente o risco, reconhecendo que ele muda conforme fontes externas, tecnologia e práticas de uso. A lição central é simples: proteção estatística não é apenas um problema de anonimização. Trata-se também de uma política

sobre o que a sociedade poderá conhecer, quais populações permanecerão estatisticamente visíveis e quem terá autoridade para decidir isso. A robustez de um sistema estatístico reside menos na adoção de uma técnica específica do que na existência de salvaguardas procedimentais que impeçam mudanças silenciosas de método ou de universo e tornem qualquer escolha auditável por aqueles que dependem dos dados.

Agradecimento

O autor agradece ao apoio recebido do projeto “Núcleo de Inteligência Artificial Responsável para Saúde – NIARSaúde UFMG”, conduzido pelo Departamento de Ciências da Computação da UFMG.

Declaração de uso de ferramentas de inteligência artificial

Durante a preparação e produção deste manuscrito, o autor utilizou o Claude Anthropic Opus 5 e Chat GPT para revisão gramatical do texto. Após o uso da ferramenta, o autor revisou, editou criticamente o conteúdo e assume total responsabilidade pela precisão e integridade do artigo final.

Referências

- ABDU, A. A.; CHAMBERS, L. M.; MULLIGAN, D. K.; JACOBS, A. Z. Algorithmic transparency and participation through the handoff lens: lessons learned from the U.S. Census Bureau’s adoption of differential privacy. *In: THE 2024 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY – FACCT. Proceedings* [...]. Rio de Janeiro, 2024.
- ABOWD, J. M.; ASHMEAD, R.; CUMINGS-MENOS, R.; GARFINKEL, S.; HEINECK, M.; HEISS, C.; JOHNS, R.; KIFER, D.; LECLERC, P.; MACHANAVAJJHALA, A.; MORAN, B.; SEXTON, W.; SPENCE, M.; ZHURAVLEV, P. The 2020 Census disclosure avoidance system topdown algorithm. *Harvard Data Science Review*, v. 2, special issue, 2022.
- ABOWD, J. M. *Third declaration of John M. Abowd*. Alabama: U.S. Department of Commerce, 2021.
- ABOWD, J. M.; SCHMUTTE, I. M. An economic analysis of privacy protection and statistical accuracy as social choices. *American Economic Review*, v. 109, n. 1, p. 171-202, 2019.
- APDU – Association of Public Data Users. *Comentário sobre a proposed rule do Census Bureau: critérios de residência e restrições a quesitos demográficos no censo decenal*. Setembro 2026.
- DALENIUS, T.; REISS, S. P. Data-swapping: a technique for disclosure control. *Journal of Statistical Planning and Inference*, v. 6, n. 1, p. 73-85, 1982.
- DINUR, I.; NISSIM, K. Revealing information while preserving privacy. *In: THE TWENTY-SECOND ACM SIGMOD-SIGACT-SIGART SYMPOSIUM ON PRINCIPLES OF DATABASE SYSTEMS. Proceedings* [...]. San Diego: Association for Computing Machinery (ACM), 2003. p. 202-210.
- HOTZ, V. J. *et al.* Balancing data privacy and usability in the federal statistical system. *PNAS*, v. 119, n. 31. 2022.

IBGE – Instituto Brasileiro de Geografia e Estatística. **Censo Demográfico 2022**. Notas metodológicas 03/2026: controle estatístico de confidencialidade dos microdados da amostra. Rio de Janeiro, 2026.

KENNY, C. T.; MCCARTAN, C.; KURIWAKI, S.; SIMKO, T.; IMAI, K. Evaluating bias and noise induced by the U.S. Census Bureau’s privacy protection methods. **Science Advances**, v. 10, n. 18, eadl2524, 2024.

MUELLER, J. T.; SANTOS-LOZADA, A. R. The 2020 US Census differential privacy method introduces disproportionate discrepancies for rural and non-white populations. **Population Research and Policy Review**, v. 41, n. 4, p. 1417-1430, 2022.

MURALIDHAR, K.; DOMINGO-FERRER, J.; SÁNCHEZ, D.; RUGGLES, S. Protecting vulnerable respondents: a critical analysis of the privacy-preserving methods of the 2010 and 2020 Decennial Census. **Population Research and Policy Review**, v. 44, n. 1, art. 8, 2025.

OFFICE FOR NATIONAL STATISTICS. Intruder testing on the 2021 Census. *In*: UNECE EXPERT MEETING ON STATISTICAL DATA CONFIDENTIALITY. **Proceedings [...]**. Wiesbaden, Germany: UNECE, 2023.

RUGGLES, S.; MAGNUSON, D. L. “It’s none of their damn business”: privacy and disclosure control in the U.S. Census, 1790-2020. **Population and Development Review**, v. 49, n. 3, 2023.

RUGGLES, S. When privacy protection goes wrong. **Journal of Economic Perspectives**, v. 38, n. 2, 2024.

RUGGLES, S. The shortcomings of synthetic census microdata. **PNAS**, v. 122, n. 11, 2025.

SU, B.; SU, W. J.; WANG, C. The 2020 US Decennial Census is more private than you (might) think. **PNAS**, v. 122, n. 45, e2500337122, 2025.

U.S. CENSUS BUREAU. Decennial census of the population of Americans; proposed residence criteria and proposed regulations for demographic questions. **Federal Register**, v. 91, n. 174, p. 57524-57531, 2026.

U.S. DEPARTMENT OF COMMERCE. **DAO 216-26** Disclosure avoidance for statistical products. 2026.

U.S. DEPARTMENT OF COMMERCE. Census Bureau. “Proposed Rules”. **Federal Register**, v. 91, n. 174, September 10, 2026.

WILLENBORG, L.; DE WAAL, T. **Elements of statistical disclosure control**. Springer, 2001.

Sobre o autor

Eduardo L. G. Rios-Neto é PhD em Demografia pela Universidade da Califórnia/Berkeley e pós-doutor pela Universidade do Texas-Austin. Professor de Demografia, aposentado e voluntário, do Centro de Desenvolvimento e Planejamento Regional (Cedeplar), Universidade Federal de Minas Gerais (UFMG). Membro da Academia Mundial de Ciências (TWAS) e membro titular da Academia Brasileira de Ciências (ABC), na área de ciências sociais. Possui Diploma da Ordem Nacional do Mérito Científico, Classe de Comendador, Ciências Sociais e Humanas, conferido pelo Ministério de Ciências e Tecnologia. Pesquisador do Cedeplar/UFMG.

Endereço para correspondência

Centro de Desenvolvimento e Planejamento Regional (Cedeplar), Universidade Federal de Minas Gerais (UFMG)
Av. Pres. Antônio Carlos, 6627, Pampulha
31270-901 – Belo Horizonte-MG, Brasil

CRediT

Reconhecimentos: Não aplicável.

Financiamento: O autor é apoiado pelo projeto 31816 | MS/ICEX/DCC/NIARSAUDE com bolsa de pesquisa de servidor docente da Fundep – Fundação de Apoio da UFMG.

Conflitos de interesse: O autor atesta que não possui qualquer interesse pessoal, comercial, acadêmico, político ou financeiro que possa gerar um conflito de interesses em relação ao manuscrito.

Aprovação ética: O autor certifica que o trabalho não inclui seres humanos ou animais.

Disponibilidade de dados e material: Os conteúdos subjacentes ao texto da pesquisa estão contidos no manuscrito.

O autor é responsável pela conceitualização; investigação; metodologia; visualização; escrita – rascunho original; escrita – revisão e edição.

Editores: Bernardo Lanza Queiroz, Júlia Almeida Calazans e Maria Carolina Tomás.

Recebido para publicação em 08/09/2026

Aceito para publicação em 15/09/2026