



## CONTROLADORIA-GERAL DA UNIÃO

### NOTA TÉCNICA Nº 1136/2022/CGAT/DTC/STPC

#### PROCESSO Nº 00190.104126/2022-86

INTERESSADO: INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA

#### 1. ASSUNTO

1.1. Análise dos documentos publicados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep em anexo à Nota de Esclarecimento acerca da adequação dos microdados dos censos educacionais e do ENEM.

#### 2. REFERÊNCIAS

2.1. SEI 2376617 - Nota de Esclarecimento sobre a divulgação dos microdados do Inep (disponível em <https://www.gov.br/inep/pt-br/assuntos/noticias/institucional/nota-de-esclarecimento-divulgacao-dos-microdados>).

2.2. SEI 2376634 - Site do INEP - Microdados (disponível em <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados>).

2.3. SEI 2376698 - Termo de Execução Descentralizada nº 8750 - TED 8750 (disponível em [https://download.inep.gov.br/microdados/TED\\_8750-UFMG.pdf](https://download.inep.gov.br/microdados/TED_8750-UFMG.pdf)).

2.4. SEI 2376690 - Nota Técnica nº 5/2021/CGCQTI/DEED/INEP (disponível em [https://download.inep.gov.br/microdados/nota\\_tecnica\\_5-2021\\_deed.pdf](https://download.inep.gov.br/microdados/nota_tecnica_5-2021_deed.pdf)).

2.5. SEI 2376741 - Nota Técnica nº 14/2021/CGIM/DAEB/INEP (disponível em [https://download.inep.gov.br/microdados/nota\\_tecnica\\_14-2021\\_daeb.pdf](https://download.inep.gov.br/microdados/nota_tecnica_14-2021_daeb.pdf)).

2.6. SEI 2376663 - Posicionamento Público da sociedade civil divulgado por 33 entidades (disponível em <https://anped.org.br/news/posicionamento-publico-de-entidades-sobre-exclusao-de-dados-do-censo-escolar-pelo-inep>).

2.7. SEI 2377995 - Opinião nº 6/2013 sobre reuso de Dados Abertos e Informações do Setor Público do *Article 29 Data Protection Working Party* (disponível em [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp207\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp207_en.pdf)).

#### 3. SUMÁRIO EXECUTIVO

3.1. Em 22/02/2022, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep publicou Nota de Esclarecimento (SEI 2376617) acerca da adequação dos microdados disponíveis em seu portal. Conforme informação constante do site do Inep (SEI 2376634), os formatos de apresentação do conteúdo dos arquivos dos microdados (conjunto de informações detalhadas relacionadas às pesquisas, aos exames e avaliações do Instituto) estariam sendo reestruturados para suprimir a possibilidade de identificação de pessoas, em atendimento às normas previstas na Lei n.º 13.709, de 14 de agosto de 2018 – Lei Geral de Proteção de Dados Pessoais (LGPD).

3.2. Ainda segundo o Instituto, o controle de privacidade nos **censos educacionais** do Inep foi analisado por meio de Termo de Execução Descentralizada nº 8750 (SEI 2376698) firmado entre o Instituto e a Universidade Federal de Minas Gerais (UFMG). Diante dos resultados das análises, a Diretoria de Estudos Educacionais (Deed) do Inep produziu nota técnica (SEI 2376690) sugerindo ações para solucionar o tratamento da privacidade nos microdados públicos.

3.3. Além dos censos educacionais, os demais exames, pesquisas e avaliações do Inep, divulgados por microdados, também foram apontados pelo Instituto como possuindo risco de reidentificação de pessoas. Assim, a Diretoria de Avaliação da Educação Básica (Daeb) elaborou nota

técnica (SEI 2376741) com a previsão de adequações no formato de divulgação dos microdados do Exame Nacional do Ensino Médio (ENEM).

3.4. Segundo Posicionamento Público da sociedade civil divulgado por 33 entidades (SEI 2376663), a adequação realizada pelo Inep significa imenso retrocesso “em termos de transparência da administração pública” e “impacto incomensurável em termos de avaliação educacional, que impede a elaboração de políticas públicas que respondam às necessidades da população”. Segundo a avaliação dessas entidades, haverá “enormes prejuízos às pesquisas e às Ciências da Educação, impedindo o desenvolvimento científico, econômico e social”.

3.5. A presente Nota Técnica possui o objetivo de analisar as duas notas técnicas citadas e o estudo da UFMG, utilizados como subsídio para a tomada de decisão do Inep de excluir as bases de anos anteriores referentes aos censos educacionais e ao Enem e publicá-las em formato significativamente menos detalhado.

#### **4. DA COMPETÊNCIA PARA ANÁLISE DO PRESENTE CASO**

4.1. A publicação de bases de dados pela Administração Pública deve atender ao interesse público, determinações legais ou ao direito de acesso à informação. Em todas as hipóteses, a divulgação motivada e proativa dessas bases de dados configura o que se denomina transparência ativa.

4.2. A Lei nº 12.527/2011 (Lei de Acesso à Informação - LAI), estabelece a obrigatoriedade de publicação de determinados dados pelos gestores públicos, em transparência ativa, mas outras normas legais e infralegais também contemplam exigências de publicação de informações pela Administração Pública. Em que pese a existência de tais obrigações em legislação esparsa, o acesso à informação pela sociedade está fundamentado na Constituição, enquanto garantia constitucional, e regulado pela LAI.

4.3. No âmbito do Poder Executivo federal, a publicação de bases de dados está regulamentada por meio do Decreto nº 7.724/2012 e, especificamente, pelo Decreto nº 8.777/2016, o qual instituiu a Política de Dados Abertos do Poder Executivo federal. A Controladoria-Geral da União (CGU) possui a competência de monitorar a implementação da LAI, conforme o art. 68, inciso IV do Decreto nº 7.724/2012. Adicionalmente, segundo o art. 10 do Decreto nº 8.777/2016, à CGU compete monitorar a aplicação do disposto nesse decreto e o cumprimento dos prazos e procedimentos previstos na Política de Dados Abertos do Poder Executivo federal.

4.4. A CGU exerce suas competências de monitoramento da LAI e da Política de Dados Abertos no âmbito do Poder Executivo federal por meio da Secretaria de Transparência e Prevenção da Corrupção (STPC), à qual compete promover e monitorar o cumprimento do disposto nos art. 68 e art. 69 do Decreto nº 7.724/2012, em articulação com as demais unidades da CGU, conforme art. 16, inciso VI, do Decreto nº 9.681/2019, que aprovou a estrutura regimental da CGU, dentre outras disposições.

4.5. Com efeito, no exercício das competências de monitoramento da implementação da LAI e da aplicação do decreto que instituiu a Política de Dados Abertos do Poder Executivo federal, a presente análise foi realizada no âmbito da STPC, com o objetivo de apresentar o entendimento da Secretaria sobre o caso concreto mencionado e de propor encaminhamentos ao Inep para a adoção de providências cabíveis, em face das conclusões consignadas ao final desta Nota Técnica.

#### **5. ARCABOUÇO LEGAL**

5.1. A LAI é um marco legal reconhecido internacionalmente como um divisor de águas. O Art. 3º da lei mudou a cultura do Estado brasileiro até então existente, ao definir em seu inciso I a observância da publicidade como preceito geral e do sigilo como exceção. De maneira harmônica, previu, ainda, que o tratamento das informações pessoais deve ser feito de forma transparente e com respeito à intimidade, vida privada, honra e imagem das pessoas, bem como às liberdades e garantias individuais, restringindo, como regra, o acesso a esse tipo de informação pessoal, exceto nas hipóteses previstas no inciso II do §1º e no §3º do seu Art. 31. Assim, desde sua vigência em 2012, adota-se no Brasil a publicidade como regra, sem descuidar-se da proteção dos dados sigilosos e pessoais. O Art. 34 da LAI é muito claro ao prever que órgãos e entidades públicas respondem diretamente pelos danos causados em decorrência da divulgação não autorizada ou utilização indevida de informações sigilosas ou informações pessoais, cabendo a apuração de responsabilidade funcional nos casos de dolo ou culpa.

5.2. A Lei nº 13.709/2018 (Lei Geral de Proteção de Dados Pessoais - LGPD) somou-se ao arcabouço jurídico até então vigente, não ao inovar quanto à obrigação de proteger os dados pessoais contra divulgação não autorizada ou indevida, mas sim ao detalhar os métodos e meios de tratamento das

informações pessoais. Assim, da mesma forma como o Art. 3º e o Art. 31 da LAI compatibilizaram no mesmo diploma legal a publicidade como preceito geral e a necessidade de proteção dos dados sigilosos e pessoais, a LGPD compatibiliza-se com a LAI e deve, com esta, ser interpretada de forma harmônica.

5.3. Exatamente nessa linha, o Enunciado nº 4/2022 da Controladoria-Geral da União (CGU) destaca que a LAI, a LGPD e a Lei nº 14.129/2021 (Lei de Governo Digital) são sistematicamente compatíveis entre si e harmonizam os direitos fundamentais do acesso à informação, da intimidade e da proteção aos dados pessoais, não havendo antinomia entre seus dispositivos:

*Nos pedidos de acesso à informação e respectivo recursos, as decisões que tratam da publicidade de dados de pessoas naturais devem ser fundamentadas nos arts. 3º e 31 da Lei nº 12.527/2011 (Lei de Acesso à Informação - LAI), vez que:*

*A LAI, por ser mais específica, é a norma de regência processual e material a ser aplicada no processamento desta espécie de processo administrativo; e*

*A LAI, a Lei nº 14.129/2021 (Lei de Governo Digital) e a Lei nº 13.709/2018 (Lei Geral de Proteção de Dados Pessoais - LGPD) são sistematicamente compatíveis entre si e harmonizam os direitos fundamentais do acesso à informação, da intimidade e da proteção aos dados pessoais, não havendo antinomia entre seus dispositivos.*

5.4. Esclarecida a inexistência de conflito entre a LAI e a LGPD, passa-se à contextualização dos avanços, também reconhecidos internacionalmente, na publicação de Dados Abertos Governamentais no Brasil.

5.5. O Brasil possui compromisso de longa data com a abertura de dados governamentais. De acordo com a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), o país conta com um sólido arcabouço legal para estruturar sua agenda de Dados Abertos Governamentais. Segundo avaliação da OCDE realizada entre os anos de 2020 e 2022, o Brasil possui, ainda, mecanismos e iniciativas específicas de governança que ajudam a alavancar e monitorar a implementação de dados governamentais abertos no país. A própria LAI possui comandos relacionados diretamente com o direito de se obter informação contida em registros produzidos ou acumulados por seus órgãos ou entidades (Art. 7º, inciso II) e com a divulgação, independentemente de requerimentos, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas (Art. 8º). O governo brasileiro é também um dos oito países fundadores da Parceria de Governo Aberto (OGP), cujas iniciativas nos planos de ação nacionais têm sido intimamente ligadas à transparência. Todos esses esforços são refletidos no desempenho do Brasil acima das médias dos países da OCDE e dos países da América Latina, de acordo com o *OURdata Index (The OECD Open, Useful Re-usable Government Data (OURdata) Index)*.

5.6. De forma a materializar os compromissos de Governo Aberto constantes dos planos de ação nacionais da OGP, o governo brasileiro criou a Política de Dados Abertos do Poder Executivo federal que define regras para disponibilização de dados abertos governamentais no âmbito do Poder Executivo federal. A Política é constituída por uma série de normativos, dentre os quais destacam-se o Decreto nº 8.777/2016, o Decreto nº 9.903/2019 e a Resolução nº 3 do Comitê Gestor da INDA (Infraestrutura Nacional de Dados Abertos). O órgão responsável pela coordenação da gestão da Política é a CGU, por meio da Infraestrutura Nacional de Dados Abertos (Art. 5º do Decreto nº 8.777/2016). A CGU também é responsável pelo monitoramento dos dispositivos da Política de Dados Abertos do Poder Executivo federal (Art. 10 do Decreto nº 8.777/2016).

5.7. Dentre os principais objetivos da Política de Dados Abertos do Poder Executivo federal estão (Art. 1º do Decreto nº 8.777/2016):

- promover a publicação de dados contidos em bases de dados de órgãos e entidades da administração pública federal autárquica e fundacional sob a forma de dados abertos;
- aprimorar a cultura de transparência pública;
- franquear aos cidadãos o acesso, de forma aberta, aos dados produzidos ou acumulados pelo Poder Executivo Federal, sobre os quais não recaia vedação expressa de acesso; e
- fomentar o controle social e o desenvolvimento de novas tecnologias destinadas à construção de ambiente de gestão pública participativa e democrática e à melhor oferta de serviços públicos para o cidadão;
- fomentar a pesquisa científica de base empírica sobre a gestão pública.

5.8. Importante ressaltar que o próprio Decreto nº 8.777/2016 estabelece no Art. 5º, §3º, que “a INDA poderá estabelecer normas complementares relacionadas com a elaboração do Plano de Dados

Abertos, bem como relacionadas a proteção de informações pessoais na publicação de bases de dados abertos nos termos deste Decreto”.

5.9. Assim, novamente percebe-se previsão normativa no âmbito da política de transparência, em especial da Política de Dados Abertos do Poder Executivo federal, para se regulamentar de forma complementar e sistemática a previsão legal do Art. 31 da LAI no que tange à publicação de bases de dados.

5.10. Posteriormente, quando a LGPD passou a integrar o ordenamento jurídico nacional, importantes detalhamentos dos métodos e meios de tratamento das informações pessoais passaram a fazer parte do arcabouço metodológico a ser interpretado em cada caso concreto de aplicação.

5.11. Ao dispor sobre o tratamento de dados pessoais, a própria LGPD apresenta a previsão de que quando uma base de dados estiver anonimizada, os “dados anonimizados não serão considerados dados pessoais para os fins desta Lei” (Art. 5º, inciso XI c/c Art. 12):

Art. 5º Para os fins desta Lei, considera-se:

(...)

**anonimização:** utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais **um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo**

(...)

Art. 12. Os **dados anonimizados não serão considerados dados pessoais para os fins desta Lei, salvo quando o processo de anonimização ao qual foram submetidos for revertido**, utilizando exclusivamente **meios próprios**, ou quando, com **esforços razoáveis, puder ser revertido**.

§ 1º A determinação do que seja razoável deve levar em consideração fatores objetivos, tais como custo e tempo necessários para reverter o processo de anonimização, de acordo com as tecnologias disponíveis, e a utilização exclusiva de meios próprios.

(...)

§ 3º A autoridade nacional poderá dispor sobre padrões e técnicas utilizados em processos de anonimização e realizar verificações acerca de sua segurança, ouvido o Conselho Nacional de Proteção de Dados Pessoais.

5.12. A discussão quanto à infalibilidade dos métodos de anonimização é ampla, havendo consenso de quanto mais tempo e recurso computacional for dedicado no processo de reidentificação, maior a probabilidade de que o processo de anonimização sobre uma base de dados possa ser revertido. Esse foi o motivo que levou a LGPD, assim como a sua equivalente legislação europeia, a prever o critério de razoabilidade na potencial reversão. Assim, a LGPD, tendo ciência de que, em teoria, qualquer base de dados anonimizada pode ser, no futuro, desanonimizada, restringiu a qualidade de reversão da anonimização para os casos em que os esforços para tal sejam razoáveis.

5.13. Conforme posicionamentos que vem sendo adotados pela Agência Nacional de Proteção de Dados (ANPD), é provável que o Brasil aplique, em definitivo, a mesma interpretação da União Europeia com relação à GDPR (General Data Protection Regulation). Nos termos da Opinião nº 6/2013 sobre reuso de Dados Abertos e Informações do Setor Público do *Article 29 Data Protection Working Party* (SEI 2377995):

Os controladores devem avaliar se um indivíduo pode ser razoavelmente identificado a partir do conjunto de dados 'anonimizado' a ser disponibilizado para reutilização e de outros dados. Em outras palavras, se qualquer organização ou indivíduo poderia identificar qualquer indivíduo a partir dos dados que estão sendo liberados – em si ou em combinação com outras informações disponíveis.

5.14. No caso dessa interpretação prevalecer, bastaria que fosse possível um único indivíduo ter seus dados reidentificados (inclusive por ele próprio) na base de dados para que a base não seja considerada anonimizada.

5.15. A partir da constatação da existência de dados pessoais na base de dados (seja identificando diretamente os titulares ou permitindo que, de forma indireta, pelo menos um titular seja reidentificado), deve ser realizada análise por parte do gestor público que resultará na decisão quanto à publicação ou não de dados de pessoas naturais fundamentada nos arts. 3º e 31 da Lei nº 12.527/2011 (Lei de Acesso à Informação - LAI), nos termos do Enunciado nº 4/2022 da Controladoria-Geral da União (CGU):

Art. 3º Os procedimentos previstos nesta Lei destinam-se a assegurar o direito fundamental de acesso à informação e devem ser executados em conformidade com os princípios básicos da administração pública e com as seguintes diretrizes:

- I - observância da publicidade como preceito geral e do sigilo como exceção;
- II - divulgação de informações de interesse público, independentemente de solicitações;
- III - utilização de meios de comunicação viabilizados pela tecnologia da informação;
- IV - fomento ao desenvolvimento da cultura de transparência na administração pública;
- V - desenvolvimento do controle social da administração pública.

(...)

Art. 31. **O tratamento das informações pessoais** deve ser feito de forma transparente e com respeito à **intimidade, vida privada, honra e imagem das pessoas**, bem como às **liberdades e garantias individuais**.

§ 1º As informações pessoais, a que se refere este artigo, relativas à intimidade, vida privada, honra e imagem:

I - terão seu **acesso restrito**, independentemente de classificação de sigilo e pelo prazo máximo de 100 (cem) anos a contar da sua data de produção, a agentes públicos legalmente autorizados e à pessoa a que elas se referirem; e

II - **poderão ter autorizada sua divulgação ou acesso por terceiros** diante de **previsão legal** ou **consentimento expresso da pessoa a que elas se referirem**.

§ 2º Aquele que obtiver acesso às informações de que trata este artigo será responsabilizado por seu uso indevido.

§ 3º **O consentimento referido no inciso II do § 1º não será exigido quando as informações forem necessárias:**

I - à prevenção e diagnóstico médico, quando a pessoa estiver física ou legalmente incapaz, e para utilização única e exclusivamente para o tratamento médico;

II - à realização de estatísticas e pesquisas científicas de evidente interesse público ou geral, previstos em lei, sendo vedada a identificação da pessoa a que as informações se referirem;

III - ao cumprimento de ordem judicial;

IV - à defesa de direitos humanos; ou

V - **à proteção do interesse público e geral preponderante**.

§ 4º A restrição de acesso à informação relativa à vida privada, honra e imagem de pessoa não poderá ser invocada com o intuito de prejudicar processo de apuração de irregularidades em que o titular das informações estiver envolvido, bem como em ações voltadas para a recuperação de fatos históricos de maior relevância.

§ 5º Regulamento disporá sobre os procedimentos para tratamento de informação pessoal.

5.16. Assim, **o fato de uma base de dados não estar anonimizada não pode ser usado como fundamento para que os dados pessoais nela contidos não se tornem públicos**, pois a LAI estabelece a publicidade como preceito geral (Art. 3º, I), determina que a divulgação de informações de interesse público deve ocorrer independentemente de solicitações (Art. 3º, II, e Art. 8º), e prevê hipóteses nas quais a publicação de dados pessoais é autorizada, das quais destaca-se:

1. Quando houver **previsão legal** (Art. 31, §1º, inciso II).
2. Quando as informações forem necessárias à **proteção do interesse público e geral preponderante** (Art. 31, §1º, inciso II, §3º, inciso V).

5.17. Enquanto a primeira hipótese se refere à existência de lei específica que determine a publicação dos dados pessoais, a segunda está relacionada à **preponderância da proteção do interesse público e geral** sobre a **proteção dos dados pessoais do titular**, em ponderação entre as respectivas garantias constitucionais subjacentes.

5.18. Não por acaso, ambas as possibilidades encontram equivalência na LGPD quando o Art. 7º traz as hipóteses nas quais o tratamento de dados pessoais poderá ocorrer:

Art. 7º O tratamento de dados pessoais somente poderá ser realizado nas seguintes hipóteses:

(...)

II - para o cumprimento de obrigação legal ou regulatória pelo controlador;

III - pela administração pública, para o tratamento e uso compartilhado de dados necessários à execução de políticas públicas previstas em leis e regulamentos ou respaldadas em contratos, convênios ou instrumentos congêneres, observadas as disposições do Capítulo IV desta Lei;

(...)

§ 3º O tratamento de dados pessoais cujo acesso é público deve considerar a finalidade, a boa-fé e o interesse público que justificaram sua disponibilização.

5.19. Assim, seja por uma obrigação legal (Art. 31, §1º, inciso II da LAI c/c Art. 7º, inciso II e Art. 11, inciso II, alínea “a” da LGPD) ou no âmbito da execução de políticas públicas previstas em lei que sinalizem ser a proteção do interesse público e geral preponderante em face da proteção dos dados pessoais do titular (Art. 31, §1º, inciso II, §3º, inciso V da LAI c/c Art. 7º, inciso III e Art. 11, inciso II, alínea “b” da LGPD), a fundamentação pela publicação dos dados pessoais deve considerar a finalidade, a boa-fé e o interesse público que justificaram sua disponibilização (Art. 7º, §3º da LGPD).

## **6. ANÁLISE DOS DOCUMENTOS RELACIONADOS À READEQUAÇÃO DOS DADOS DOS CENSOS EDUCACIONAIS**

6.1. A Nota Técnica nº 5/2021/CGCQTI/DEED (SEI 2376690), da Diretoria de Estatísticas Educacionais (DEED) do Inep, apresentou os resultados do TED 8750 (SEI 2376698), executado pela Universidade Federal de Minas Gerais (UFMG) e recomendou as medidas a serem adotadas com relação à readequação dos formatos dos microdados dos censos educacionais (Censo da Educação Superior e Censo da Educação Básica) nos termos propostos pelo estudo resultado do TED 8750.

6.2. Segundo a referida Nota Técnica, “esses resultados obtidos [conclusões do estudo realizado pela UFMG] demonstram de forma inequívoca que a atual forma de divulgação dos Censos Educacionais pelo Inep submete os titulares dos dados a consideráveis riscos de violação de privacidade, incluindo reidentificação”, o que “poderia constituir em violação da Lei Geral de Proteção de Dados Pessoais”.

6.3. O estudo da UFMG realizou análises quantitativas utilizando diferentes modelos e tipos de ataque. Os resultados quantitativos foram obtidos como resultado do tipo “Ataques de reidentificação” (página 14 do Produto 02 do TED 8750 - 2376698) com o objetivo de “determinar uma correspondência entre registros presentes na base de dados e os titulares desses registros”, visando avaliar a possibilidade de reidentificação nos termos do caput do Art. 12 da LGPD: “salvo quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido”.

6.4. Em resumo, a UFMG concluiu que é possível reidentificar 38,87% dos alunos constantes dos microdados do Censo da Educação Superior de 2018, quando 3 quaseidentificadores (dia e ano do nascimento; e código do curso) estiverem disponíveis, permitindo o cruzamento de informações e a identificação do titular dos dados pessoais. Esse percentual aumenta para 79,20% quando 4 quaseidentificadores estiverem disponíveis (dia, mês e ano do nascimento; e código do curso), chegando até 97,22% dos alunos reidentificados, quando todos os 11 quaseidentificadores estiverem disponíveis (dia, mês e ano do nascimento; código do curso; gênero; cor/raça; código do município do nascimento; nacionalidade; código do país de origem; código da IES; código do curso; e escola de conclusão do ensino médio), conforme detalhado na coluna “Sucesso determinístico” da tabela 4.3.4 do Experimento 8 (página 40 do Produto 02 do TED 8750 - 2376698).

6.5. De maneira análoga, a UFMG concluiu que é possível reidentificar 14,54% dos alunos constantes dos microdados do Censo da Educação Básica de 2018, quando 3 quaseidentificadores (ano do nascimento; código do município do nascimento; e código da entidade) estiverem disponíveis, permitindo o cruzamento de informações e a identificação do titular dos dados pessoais. Esse percentual aumenta para 33,12% quando 4 quaseidentificadores estiverem disponíveis (mês e ano do nascimento; código do município do nascimento; e código da entidade), chegando a no máximo 60,90% dos alunos reidentificados, quando todos os 10 quaseidentificadores estiverem disponíveis (mês e ano do nascimento; gênero; cor/raça; nacionalidade; código do país de origem; código do município do nascimento; código do município de residência; código da Entidade; Dependência Administrativa: Federal, Estadual, Municipal ou Privada), conforme detalhado na coluna “Sucesso determinístico” da tabela 4.3.7 do Experimento 9 (página 43 do Produto 02 do TED 8750).

6.6. A conclusão de que 97,22% dos alunos são reidentificáveis na base de microdados do Censo da Educação Superior de 2018 possui como premissa a obtenção como informação auxiliar dos valores dos quaseidentificadores dos 10.511.038 alunos (97,22% dos 10.811.601 alunos únicos da base do Censo da Educação Superior), contendo, para cada aluno, os seguintes campos: dia, mês e ano do nascimento; código do curso; gênero; cor/raça; código do município do nascimento; nacionalidade; código do país de origem; código da IES; código do curso; e escola de conclusão do ensino médio.

6.7. De forma análoga, a conclusão de que 60,90% dos alunos são reidentificáveis na base de microdados do Censo da Educação Básica de 2018 possui como premissa a obtenção como informação auxiliar dos valores dos quaseidentificadores dos 29.013.261 alunos (60,90% dos 47.640.822 alunos únicos da base do Censo da Educação Básica), contendo, para cada aluno, os seguintes campos: mês e ano do nascimento; gênero; cor/raça; nacionalidade; código do país de origem; código do município do nascimento; código do município de residência; código da Entidade; Dependência Administrativa: Federal, Estadual, Municipal ou Privada.

6.8. Embora seja necessária a obtenção dessas bases de dados hipotéticas para que se possa afirmar quais seriam os percentuais exatos de alunos que seriam reidentificados com "absoluta certeza", fato é que a possibilidade de reidentificação de um único indivíduo já impede que a base de dados seja considerada anonimizada. Assim, qualquer base que possua algum registro cujos atributos não são partilhados com pelo menos um outro registro (k-anonimato com  $k=1$ ) não poderia ser considerada anonimizada.

6.9. Nesse sentido, já se poderia afirmar que ambas as bases dos censos educacionais não estariam anonimizadas, contendo, assim, dados pessoais que poderiam de forma indireta levar a reidentificação de parte de seus titulares. Nessa linha, o próximo passo seria a avaliação quanto à existência de hipóteses no âmbito do Art. 31 da LAI que autorizariam a publicação desses dados pessoais, conforme discutido nos itens 5.15 a 5.19 desta Nota Técnica.

6.10. Nessa linha, embora o Parecer nº 00018/2022/PROC/PFINEP/PGF/AGU não faça citação explícita aos Art. 3º e 31 da LAI, que é a norma regente para a tomada de decisão quanto à publicação de bases de dados mesmo quando contiver dados pessoais sensíveis, o documento em seu item 19 apresenta de forma clara os dispositivos da LGPD que são sistematicamente compatíveis com a LAI:

19. Nesse sentido, no caso de execução de políticas públicas, enfatiza-se que se faz necessário demonstrar que a política pública está inserida nas atribuições legais do órgão e que a política pública está prevista "...em leis e regulamentos ou respaldadas em contratos, convênios ou instrumentos congêneres", no caso de dados pessoais (art. 7º, III, Lei nº 13.709/2018) ou "... em leis ou regulamentos", no caso de dados pessoais sensíveis (art. 11, II, "b", Lei nº 13.709/2018).

6.11. O Parecer traz, ainda, referência ao Anexo II do "Guia sobre Tratamento de Dados Pessoais pelo Poder Público", publicado pela ANPD, que apresenta alguns parâmetros e recomendações relacionados com "Cuidados a serem observados quando da divulgação de dados pessoais pelo Poder Público", muito relacionados com a preocupação em se considerar a finalidade, a boa-fé e o interesse público que justificaram a disponibilização da informação, disposta no Art. 7º, §3º da LGPD.

6.12. Assim, observa-se que o Parecer cita no seu item 19 dispositivos apresentados no item 5.19 desta Nota Técnica quando da fundamentação pela publicação dos dados pessoais (inclusive sensíveis) no âmbito da execução de políticas públicas previstas em lei que sinalizem ser a proteção do interesse público e geral preponderante à proteção dos dados pessoais do titular.

6.13. Entretanto, como não foram tornados públicos outros documentos que eventualmente a área jurídica do Inep possa ter tido acesso, não foi possível identificar o motivo pelo qual o Parecer conclui pela impossibilidade da divulgação das bases não anonimizadas:

24. Portanto, respondendo aos termos da consulta formulada no OFÍCIO Nº 0855589/2022/AGGE/GABINEP (0855589), e também em atenção às disposições do inciso IV do art. 4º e inciso III do art. 6 da Lei de Acesso à Informação - LAI, Lei nº 12.527/2011, caso se vislumbre - como no estudo resultante do Termo de Execução Descentralizada (TED) 8750, que mesmo diante da remoção de identificadores (desidentificação) ou de sua substituição por código artificialmente criado (pseudonimização), seja possível realizar a reidentificação ou inferir atributos sensíveis dos indivíduos titulares dos microdados constantes das bases mantidas nos censos educacionais, deve o INEP restringir o acesso para que se evite a obtenção de dados pessoais, sujeitando o amplo acesso e divulgação à permissão dos titulares dos dados ou concedendo de acordo com os objetivos de pesquisa ou do interessado, desde que em salas seguras, conforme diagnóstico preconizado pela Diretoria de Estatísticas Educacionais em consideração aos resultados do TED 8750.

25. Nessa linha de entendimento, (...), cabe ressaltar que - se a divulgação dos censos ou outras bases de dados mantidos pelo INEP puder resultar em acesso, por terceiros, a microdados pessoais não anonimizados ou que permitam a reidentificação de seus titulares, a divulgação não poderá ser realizada, de acordo com a LGPD.

6.14. Diante dos documentos analisados, não foi possível identificar a fundamentação utilizada pelo gestor federal para motivar a mudança no formato de disponibilização dos microdados dos censos educacionais.

6.15. É essencial destacar que esta Coordenação-Geral considera que o mero fato de a base de dados não estar anonimizada não pode ser usado como fundamento para que os dados pessoais nela contidos não se tornem públicos, pois a LAI estabelece a publicidade como preceito geral (Art. 3º, I), determina que a divulgação de informações de interesse público deve ocorrer independentemente de solicitações (Art. 3º, II, e Art. 8º), e prevê hipóteses nas quais a publicação de dados pessoais é autorizada, dentre as quais previsão legal (Art. 31, §1º, inciso II) e a proteção do interesse público e geral preponderante (Art. 31, §1º, inciso II, §3º, inciso V). Na mesma linha, verifica-se que a anonimização dos dados não é considerada pela LGPD como uma condição para o tratamento de dados com a finalidade de publicação pela Administração Pública, motivo pelo qual não se trata da única medida possível de tratamento dos dados no caso concreto analisado. Por fim, o consentimento dos titulares dos dados não é exigido no caso em tela, quando as informações forem necessárias à proteção do interesse público e geral preponderante, nos termos do Art. 31 da LAI.

## 7. ANÁLISE DOS DOCUMENTOS RELACIONADOS À READEQUAÇÃO DOS DADOS DO ENEM

7.1. A Nota Técnica nº 14/2021/CGIM/DAEB, de 28/07/2021, foi elaborada pela Diretoria de Avaliação da Educação Básica (DAEB) do Inep com o objetivo de “fundamentar a necessidade de modificações urgentes nos microdados do Exame Nacional do Ensino Médio (ENEM) divulgados anualmente no portal do Inep, de modo a atender as exigências trazidas pela recente Lei Geral de Proteção de Dados (LGPD)”.

7.2. A Nota cita que os microdados da edição de 2020 do ENEM encontravam-se prontos para divulgação, seguindo o mesmo modelo utilizado em edições anteriores, os quais eram, até então, considerados anonimizados, “uma vez que não consta na base pública qualquer variável que permita a identificação direta do participante”.

7.3. A Nota Técnica, então, ao considerar fundamental a realização de um estudo para avaliar se os microdados produzidos pela DAEB cumprem os requisitos previstos na legislação de proteção de dados pessoais, sugere, enquanto esse estudo não for realizado, a adoção de um modelo simplificado de microdados, eliminando da base pública as variáveis que, no entendimento da DAEB, facilitariam a identificação indevida do participante. Mesmo excluindo essas variáveis, a DAEB manifesta preocupação quanto ao cumprimento da LGPD:

Entretanto, é importante destacar que a CGIM não tem condições técnicas de assegurar que as medidas propostas sejam suficientes para atender os requisitos da LGPD. Por isso, depende da aprovação desse modelo pela DAEB, pela PROJUR e pela Presidência do INEP, assegurando que, ao divulgar os microdados do ENEM 2020 nesse modelo simplificado, a equipe técnica da CGIM não estará descumprindo a LGPD.

7.4. Seguem variáveis propostas pela DAEB para serem excluídas/alteradas e respectivo resumo da fundamentação:

- Exclusão do código da escola de conclusão do ensino médio: A DAEB considera ser prudente a exclusão dessa variável dos microdados do ENEM, “uma vez que a mesma permite a identificação indevida dos participantes”. Essa conclusão advém do relato de que a DAEB recebe um número considerável de demandas encaminhadas por representantes das escolas, informando que o seu resultado no ENEM por Escola estaria errado. Segundo a Nota, essas reclamações só são possíveis porque, de alguma maneira, há a identificação dos participantes por parte dos gestores dessas instituições, possivelmente porque esses indivíduos têm acesso a informações cadastrais detalhadas de seus alunos.
- Exclusão das informações referentes aos pedidos de atendimento especializado e específico, recursos de atendimento especializado e específico para a realização da prova: Segundo a Nota Técnica, essas informações ao serem dados pessoais sensíveis dependeriam de consentimento específico para serem publicados, exceto se os microdados do ENEM fossem considerados dados anonimizados. Ocorre, que, segundo interpretação da DAEB, como o número de participantes que solicitam

atendimento especializado ou específico é muito pequeno se comparado com o número total de inscritos, essas variáveis facilitariam a identificação do participante, de modo que o dado deixaria de ser anonimizado, dependendo de consentimento do seu titular para ser divulgado. Assim, nos termos da Nota Técnica, seria mais prudente remover essas variáveis dos microdados.

- Substituição da Idade por Faixa Etária: A justificativa apresentada na Nota Técnica para substituir a informação da idade do participante no dia 31 de dezembro do ano da edição do ENEM por uma faixa etária é de que a idade, combinada com outros dados cadastrais eventualmente à disposição dos usuários dos microdados, poderia facilitar a identificação indevida do participante. Ademais, segundo a DAEB, “a experiência demonstra que a substituição da variável por uma faixa etária não prejudicaria as pesquisas”.
- Exclusão de informações referentes aos municípios de nascimento e residência do participante: Segundo a Nota Técnica, “essas informações, combinadas com outros dados cadastrais eventualmente a disposição dos usuários dos microdados, podem facilitar a identificação indevida do participante”. Adicionalmente, segundo a DAEB, “a experiência demonstra que essas variáveis possuem pouca relevância para os pesquisadores e demais usuários dos microdados, uma vez que a maior parte das pesquisas utilizam como base o município do local de prova (que será mantido nos microdados)”.

7.5. Inicialmente, reforça-se que a norma de regência processual e material a ser aplicada na tomada de decisão pelo formato da publicação da base de dados do ENEM é a LAI, cabendo ao gestor fundamentar sua decisão nos arts. 3º e 31 da referida Lei, como também dispõe o recente Enunciado nº 4/2022 da CGU.

7.6. A Nota Técnica da DAEB esclarece que não foi realizado estudo específico sobre a base do ENEM nos moldes do realizado pela UFMG para os censos educacionais. Entretanto, apresenta dúvidas sobre o fato de os microdados serem de fato dados anonimizados.

7.7. Muito embora o Inep esclareça que a dúvida surgiu pela nova interpretação que a LGPD teria dado ao conceito de base de dados anonimizada, em momento nenhum na Nota Técnica da DAEB há referência a alguma evidência que indique significativo risco de prejuízo à intimidade, vida privada, honra e imagem dos titulares que eventualmente estariam sendo indiretamente identificados na base de dados publicada. Assim, a necessária análise sobre a existência de determinação legal, de interesse público preponderante aplicável ao caso, ou do cumprimento do direito de acesso à informação não foi realizada pelo gestor, o qual fundamentou sua decisão com base na conclusão de que as bases de dados não estariam anonimizadas.

7.8. No entendimento desta Coordenação-Geral, no que tange a novas atualizações de bases de dados já publicadas, sem acréscimo de novos tipos de dados, os dados atualizados devem continuar sendo publicados no formato histórico, pois trata-se dos mesmos metadados (atributos), apenas com dados atualizados. Assim, presume-se o correto enquadramento nos termos da LAI e a boa-fé do ato de publicação no formato histórico. Isso porque não há, na documentação analisada, menção a qualquer reclamação, denúncia ou notícia de que a publicação dos dados pelo Inep, no formato historicamente utilizado, tenha causado danos aos seus titulares, ou os tenha exposto a algum risco iminente de dano.

7.9. Se, por outro lado, houvesse evidências que indicassem significativo risco de prejuízo à intimidade, vida privada, honra e imagem dos titulares, direta ou indiretamente identificáveis na base de dados publicada, aí sim haveria motivação para que o Inep promovesse nova avaliação quanto ao estrito atendimento ao interesse público na divulgação desses dados e quanto a medidas de mitigação dos riscos mencionados. Caso a avaliação demonstrasse que os riscos associados à publicação dos dados fossem desproporcionais à proteção dos dados pessoais e extrapolassem o interesse coletivo e geral dessa publicação (hipóteses previstas no Art. 31 da LAI), as bases de dados históricas deveriam, então, ser retiradas do ar e a atualização periódica suspensa até que fosse possível obter-se um formato cuja publicação atendesse estritamente a sua finalidade prevista na LAI. Alternativamente a um novo formato de dados, a análise também poderia demonstrar que medidas administrativas (controle e registro de acesso aos dados pessoais) e jurídicas (termo de responsabilidade ou de confidencialidade) poderiam ser suficientes para a mitigação dos riscos atrelados à publicação dos dados, em face do interesse público de sua divulgação.

7.10. Em resumo, com a conclusão de que as bases de dados não estão anonimizadas, caberia ao gestor do Inep realizar as análises descritas nos itens 5.15 a 5.19, dentre as quais realizar a ponderação entre o interesse público de divulgação dos dados pessoais e os riscos aos seus titulares, de modo a atender à finalidade de sua publicização, adotando, sempre que necessário, as medidas proporcionais de mitigação de riscos.

## 8. CONCLUSÃO

8.1. Após a análise realizada, apresenta-se síntese sobre os principais pontos desenvolvidos ao longo desta Nota Técnica:

8.1.1. **Legislação a ser observada para subsidiar a tomada de decisão dos gestores públicos pela publicação ou não de bases de dados:** A norma de regência processual e material a ser aplicada nas decisões que tratam da publicidade de dados de pessoas naturais é a Lei nº 12.527/2011 (Lei de Acesso à Informação - LAI). As decisões devem ser fundamentadas nos arts. 3º e 31 da LAI. Adicionalmente, a Lei nº 14.129/2021 (Lei de Governo Digital) e a Lei nº 13.709/2018 (Lei Geral de Proteção de Dados Pessoais - LGPD) são sistematicamente compatíveis entre si e harmonizam os direitos fundamentais do acesso à informação, da intimidade e da proteção aos dados pessoais, conforme reconhece o Enunciado nº 4/2022 da CGU.

8.1.2. **O fato de uma base de dados não estar anonimizada não significa que não deva ser publicada:** O fato de uma base de dados não estar anonimizada não pode ser usado como fundamento para que os dados pessoais nela contidos não se tornem públicos, pois a LAI estabelece a publicidade como preceito geral (Art. 3º, I), determina que a divulgação de informações de interesse público deve ocorrer independentemente de solicitações (Art. 3º, II, e Art. 8º), e prevê hipóteses nas quais a publicação de dados pessoais é autorizada, das quais destaca-se: (i) Quando houver previsão legal (Art. 31, §1º, inciso II); e (ii) Quando as informações forem necessárias à proteção do interesse público e geral preponderante (Art. 31, §1º, inciso II, §3º, inciso V). Enquanto a primeira hipótese se refere à existência de lei específica que determine a publicação dos dados pessoais, a segunda está relacionada à preponderância da proteção do interesse público e geral sobre a proteção dos dados pessoais do titular. Assim, nos termos da LAI, a fundamentação por uma eventual negativa pela publicação de uma base de dados que contenha dados pessoais (independentemente de ser possível identificar direta ou indiretamente seus titulares), apenas poderá ocorrer após a avaliação quanto à previsão legal e existência de interesse público e geral preponderante sobre o direito de proteção dos dados pessoais (Art. 31, §1º, inciso II, e §3º, inciso V).

8.1.3. **Bases de dados contendo dados pessoais que já estejam publicadas, mas que precisem de atualização periódica, devem continuar sendo atualizadas nos moldes históricos até que ocorra a avaliação e a devida fundamentação nos termos do Art. 31, §1º, inciso II, e §3º, inciso V, da LAI:** Em regra, os dados atualizados devem continuar sendo publicados no formato histórico, pois trata-se dos mesmos metadados (atributos), apenas com dados atualizados. Assim, presume-se o correto enquadramento nos termos da LAI e a boa-fé do ato de publicação no formato histórico. Entretanto, caso surjam evidências que indiquem significativo risco de prejuízo à intimidade, vida privada, honra e imagem dos titulares, direta ou indiretamente identificáveis na base de dados publicada, deve ser realizada nova avaliação quanto ao estrito atendimento ao interesse público na divulgação desses dados e quanto a medidas de mitigação do risco mencionado (Art. 31 da LAI, §1º, inciso II, e §3º, inciso V). Caso a avaliação demonstre que os riscos associados à publicação dos dados extrapolem o interesse coletivo e geral dessa publicação (hipóteses previstas no Art. 31 da LAI), as bases de dados históricas devem, então, ser retiradas do ar para a adoção de medidas de mitigação dos riscos envolvidos, observando-se, em todo caso, a permanência da divulgação de dados que não estejam sujeitos a proteção, nos termos da LAI. Dentre as medidas de mitigação dos riscos, podem ser adotadas medidas administrativas (controle e registro de acesso aos dados pessoais), jurídicas (termo de responsabilidade ou de confidencialidade) e técnicas (anonimização ou alteração do formato da base de dados).

## 9. ENCAMINHAMENTOS

9.1. Diante da análise realizada, propõe-se:

9.1.1. Encaminhamento desta Nota Técnica ao Inep para análise quanto à adoção das seguintes providências:

9.1.1.1. Tendo em vista que as bases dos censos educacionais nos moldes históricos já foram retiradas do ar, **realizar avaliação que considere a previsão legal e a finalidade de sua publicação, e a existência de interesse público e geral preponderante sobre o direito de proteção dos dados pessoais**

(Art. 31 da LAI, §1º, inciso II, e §3º, inciso V). Caso essa avaliação conclua pelo atendimento à previsão legal e à finalidade de sua divulgação ou pela existência de interesse público e geral preponderante sobre o direito de proteção dos dados pessoais (hipóteses previstas no Art. 31 da LAI), **publicar as bases de dados nos moldes históricos (microdados)**. Caso contrário (inexistência de atendimento à previsão legal e à finalidade de publicação, e preponderância da proteção dos dados pessoais sobre o interesse público e geral na publicação), **promover** a adoção de medidas de mitigação dos riscos envolvidos, observando-se, em todo caso, a permanência da divulgação de dados que não estejam sujeitos a proteção, nos termos da LAI.

9.1.1.2. Tendo em vista que a base do ENEM nos moldes históricos já foi retirada do ar, **realizar avaliação que considere a previsão legal e a finalidade de sua publicação, e a existência de interesse público e geral preponderante sobre o direito de proteção dos dados pessoais** (Art. 31 da LAI, §1º, inciso II, e §3º, inciso V). Caso essa avaliação conclua pelo atendimento à previsão legal e à finalidade de sua divulgação ou pela existência de interesse público e geral preponderante sobre o direito de proteção dos dados pessoais (hipóteses previstas no Art. 31 da LAI), **publicar as bases de dados nos moldes históricos (microdados)**. Caso contrário (inexistência de atendimento à previsão legal e à finalidade de publicação, e preponderância da proteção dos dados pessoais sobre o interesse público e geral na publicação), **promover** a adoção de medidas de mitigação dos riscos envolvidos, observando-se, em todo caso, a permanência da divulgação de dados que não estejam sujeitos a proteção, nos termos da LAI.

9.1.1.3. Para todas as bases de dados restantes que já estejam publicadas, **mesmo que haja dúvidas sobre sua correta anonimização**, tendo em vista a presunção do correto enquadramento nos termos da LAI e a boa-fé do ato de publicação no formato histórico, **manter todas as bases de dados publicadas, exceto se surgirem evidências que indiquem significativo risco de prejuízo à intimidade, vida privada, honra e imagem dos titulares, indiretamente identificados ou identificáveis** em alguma base de dados publicada. Nesse caso, deve ser realizada nova avaliação quanto à existência de previsão legal ou de interesse público e geral preponderante sobre o direito de proteção dos dados pessoais (Art. 31 da LAI, §1º, inciso II, e §3º, inciso V). Caso essa avaliação demonstre a existência de previsão legal para publicação dos dados pessoais ou interesse público e geral preponderante sobre o direito de proteção dos dados pessoais (hipóteses previstas no Art. 31 da LAI), **continuar publicando e a base de dados nos moldes históricos**. Caso contrário (inexistência de previsão legal e preponderância da proteção dos dados pessoais sobre o interesse público e geral na publicação), retirar do ar a base de dados, suspender a atualização periódica e promover a compatibilização das bases de dados à LAI, por exemplo utilizando técnicas que anonimem a base de dados ou que alterem a sua estrutura de forma que uma nova avaliação a enquadre nas hipóteses previstas no Art. 31 da LAI.

9.1.2. Encaminhamento desta Nota Técnica à Autoridade Nacional de Proteção de Dados – ANPD para conhecimento.

À consideração superior.

**ROGERIO VIEIRA DOS REIS**  
Coordenador-Geral de Governo Aberto e Transparência

De acordo, encaminho à consideração superior.

**BRENO BARBOSA CERQUEIRA ALVES**  
Diretor de Transparência e Controle Social

De acordo, encaminhe-se conforme proposto.

**ROBERTO CÉSAR DE OLIVEIRA VIÉGAS**  
Secretário de Transparência e Prevenção da Corrupção

---



Documento assinado eletronicamente por **ROGERIO VIEIRA DOS REIS, Coordenador-Geral de Governo Aberto e Transparência**, em 23/05/2022, às 18:12, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

---



Documento assinado eletronicamente por **BRENO BARBOSA CERQUEIRA ALVES, Diretor de Transparência e Controle Social**, em 23/05/2022, às 18:13, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

---



Documento assinado eletronicamente por **ROBERTO CESAR DE OLIVEIRA VIEGAS, Secretário de Transparência e Prevenção da Corrupção**, em 24/05/2022, às 19:26, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site <https://sei.cgu.gov.br/conferir> informando o código verificador 2376783 e o código CRC 2B74F11A