



AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS
Coordenação-Geral de Fiscalização

Nota Técnica nº 12/2023/CGF/ANPD

1. INTERESSADO

1.1. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP).

2. ASSUNTO

2.1. Avaliação dos Relatórios de Impacto à Proteção de Dados elaborados pelo INEP para fins adequação da divulgação dos microdados do censo escolar e do Enem à Lei nº 13.709, de 14 de agosto de 2018 — Lei Geral de Proteção de Dados Pessoais (LGPD).

3. REFERÊNCIAS

- 3.1. Processo SEI nº 00261.000730/2022-53;
- 3.2. Nota de Esclarecimento INEP (SEI nº 3289150);
- 3.3. Nota Técnica INEP 5/2021/CGCQTI/DEED (SEI nº 3289210);
- 3.4. Nota Técnica INEP 14/2021/CGIM/DAEB (SEI nº 3289220);
- 3.5. Termo de Execução Descentralizada UFMG (SEI nº 3289230);
- 3.6. Parecer 00018/2022/PROC/PFINEP/PGF/AGU (SEI nº 3289237);
- 3.7. Manifestação pública de entidades (SEI nº 3289249);
- 3.8. Nota Técnica nº 46/2022/CGF/ANPD (SEI nº 3319546);
- 3.9. Nota Técnica nº 136/2022/CGAT/DTC/STPC (SEI nº 3414875), da Controladoria Geral da União (CGU);
- 3.10. Relatório de Impacto à Proteção de Dados Pessoais dos Microdados do Exame Nacional do Ensino Médio (ENEM) (SEI 3848205);
- 3.11. Relatório de Impacto à Proteção de Dados Pessoais dos Microdados dos Censos da Educação (SEI 3848206).

4. RELATÓRIO^{III}

4.1. O processo em tela tem por origem a suspensão da divulgação dos microdados do censo escolar e do Exame Nacional do Ensino Médio (Enem) pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Em Nota de Esclarecimento publicada pelo INEP (SEI nº 3289150) no sítio eletrônico do Instituto, a autarquia federal informou que a adequação dos microdados disponíveis em seu portal estaria em curso com base em estudos técnicos e análises jurídicas que priorizassem o pleno atendimento às exigências previstas na Lei nº 13.709, de 14/08/2018 — Lei Geral de Proteção de Dados Pessoais (LGPD). De acordo com a nota, a necessidade de adequação das divulgações decorre da constatação de risco potencial de identificação das pessoas a quem os dados estatísticos se referem, conforme estudo realizado em parceria com a Universidade Federal de Minas Gerais (UFMG) por meio de Termo de Execução Descentralizada (TED 8750, SEI nº 3289230) e executado pelo Laboratório Inscript (Laboratory of Information Security, Cryptography, Privacy, and Transparency) do Departamento de Ciência da Computação (DCC) da universidade.

4.2. No referido estudo, foi constatado que as técnicas de proteção de privacidade utilizadas nos microdados para retirada de identificadores individuais, como nome e número do Cadastro de Pessoas Físicas (CPF) e Registro Civil, não são suficientes para assegurar a privacidade dos respondentes. O estudo conclui que a forma como então ocorria a divulgação dos censos educacionais pelo INEP submete os titulares dos dados a consideráveis riscos de violação de privacidade, incluindo reidentificação e inferência de atributos sensíveis, circunstâncias que poderiam constituir violação à LGPD.

4.3. Adicionalmente, no Parecer nº 00018/2022/PROC/PFINEP/PGF/AGU (SEI nº 3289237), a Procuradoria Federal Especializada junto ao INEP (PFE-INEP) manifestou entendimento de que a divulgação dos microdados públicos do censo escolar nos moldes anteriores infringia as disposições legais vigentes, em especial, a LGPD. Acrescentou que, de acordo com a situação apresentada à época, "não se identificam disposições legais capazes de ofertar segurança jurídica para a divulgação de dados cujo acesso possa resultar em reidentificação dos titulares dos dados, nos termos da LGPD." Por fim, concluiu que "o INEP tem o poder-dever de suspender a divulgação de dados que apresentem risco potencial de violação à LGPD, tal qual demonstrado no presente parecer e justificado tecnicamente nos resultados apresentados no Termo de Execução Descentralizada (TED) 8750".

4.4. Ao passo que suspendeu a divulgação dos microdados dos censos escolares e do Enem, amparado posteriormente pelo parecer de sua Procuradoria Federal Especializada, o INEP sugeriu um conjunto de ações para mitigar o risco à violação de privacidade dos titulares no tratamento indevido dos microdados públicos, visto que o modelo até então disponibilizado se mostrou inadequado. Em sua Nota de Esclarecimento, o Instituto ressaltou que as pesquisas que utilizam os dados tratados pelo Instituto - e que foram eventualmente restringidos no novo modelo de divulgação -, não estariam inviabilizadas, pois haveria diversos meios de acessar essas informações. Como exemplo, o INEP menciona o Serviço de Acesso a Dados Protegidos (SEDAP), que possibilitaria o uso de bases restritas por pesquisadores, desde que observados os protocolos de segurança do serviço. Esse mecanismo permitiria, segundo o INEP, o desenvolvimento de estudos amplos e detalhados, considerando tendências, padrões e trajetórias educacionais traçados a partir de evidências apuradas pela Autarquia. O INEP também se comprometeu a apresentar uma forma de expandir a sua capacidade de disponibilização de dados de forma segura nos Estados, por meio de parcerias com instituições federais de educação superior.

4.5. A remoção dos microdados do sítio eletrônico do INEP ocasionou repercussão pública, conforme se pode observar no documento Manifestação Pública de Entidades (SEI nº 3289249). Nesse documento, as entidades defenderam que é preciso proteger a privacidade, sem abdicar da transparência, razão pela qual a LGPD não poderia ser utilizada como justificativa genérica para o descarte dos microdados, por carecer de fundamento legal. As entidades alegaram, ainda, que a própria LGPD, art. 7º, II e III, autoriza que a administração pública realize o tratamento de dados pessoais necessários ao cumprimento de obrigação legal e/ou execução de políticas públicas, sem que para isso seja necessário o prévio consentimento do titular destes dados. Nessa senda, o tratamento de dados pessoais durante o censo escolar pelo INEP estaria baseada no

cumprimento da própria Constituição Federal que determina, em seu art. 208, §3º, que o Poder Público deverá recensear educandos no ensino fundamental, bem como na Resolução nº 1 de 2018 do CNE/MEC, que institui as diretrizes operacionais para a coleta e registro de dados cadastrais de estudantes e profissionais de educação que atuam em instituições públicas e privadas de ensino em todo o território nacional.

4.6. A Controladoria Geral da União (CGU) também se manifestou acerca dos fatos por meio da Nota Técnica nº 136/2022/CGAT/DTC/STPC (SEI nº 3414875). No documento, a CGU destaca o compromisso do governo brasileiro com a transparência e abertura de dados governamentais. Defende que, 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. A CGU sugere, ainda, “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 de dados pessoais (art. 31 da LAI, §1º, inciso II, e §3º, inciso V)”.

4.7. Após a análise dos documentos mencionados acima, a CGF elaborou a Nota Técnica nº 46/2022/CGF/ANPD (SEI nº 3319546), encaminhada ao INEP para ciência e providência de elaboração do Relatório de Impacto à Proteção de Dados (RIPD), para fins adequação da divulgação dos microdados do censo escolar e do Enem à LGPD. Nesse sentido, a ANPD determinou que o INEP deveria elaborar o RIPD, em observância aos princípios previstos na LGPD, com o intuito de avaliar os riscos a que são submetidos os titulares com a divulgação dos microdados. Na visão da CGF, após a elaboração do RIPD, o INEP terá condições de decidir sobre a amplitude da divulgação dos dados, sendo possível a disponibilização dos microdados em versões diferentes para a sociedade e para instituições de pesquisa, mediante termo de responsabilidade.

4.8. Em 29.12.2022, o INEP apresentou à ANPD o Relatório de Impacto à Proteção de Dados Pessoais dos Microdados do Exame Nacional do Ensino Médio (ENEM) (SEI 3848205), bem como o Relatório de Impacto à Proteção de Dados Pessoais dos Microdados dos Censos da Educação (SEI 3848206).

4.9. É o relatório.

5. ANÁLISE

I. COMPETÊNCIA DA ANPD NO PRESENTE CASO

5.1. A controvérsia acerca da metodologia de divulgação dos microdados dos censos escolares e do Enem face à LGPD, tendo em vista o risco que ela representa aos direitos dos titulares de dados, bem como a repercussão pública e os desdobramentos do caso indicam a necessidade de manifestação desta autoridade. A ANPD é o órgão que detém a competência de (i) zelar pela proteção de dados pessoais; (ii) fiscalizar e aplicar sanções em caso de tratamento de dados em descumprimento à legislação; (iii) promover na população o conhecimento das normas e das políticas públicas sobre proteção de dados pessoais e das medidas de segurança; (iv) ouvir os agentes de tratamento e a sociedade em matérias de interesse relevante e prestar contas sobre suas atividades e planejamento e (v) deliberar, na esfera administrativa, em caráter terminativo, sobre a interpretação da LGPD, as suas competências e os casos omissos, conforme previsão do art. 55-J, incisos I, IV, VI, XIV, XX da LGPD.

5.2. Em sua Nota Técnica nº 46, esta Coordenação-Geral de Fiscalização - CGF expôs seu entendimento sobre o caso concreto e forneceu orientações gerais para a publicização dos microdados, em respeito à privacidade dos titulares e sem prejuízo à transparência das informações. Por fim, solicitou como providência a elaboração do Relatório de Impacto à Proteção de Dados para fins adequação da divulgação dos microdados do censo escolar e do Enem à LGPD.

5.3. A presente Nota Técnica tem por objetivo retomar as principais questões envolvidas no processo em tela, reforçando o entendimento da CGF no que tange à sua esfera de competência, bem como avaliar os relatórios apresentados pelo INEP quanto a sua suficiência e adequação para mitigar os riscos aos direitos dos titulares de dados identificados ao longo do processo.

II. O INEP E A DIVULGAÇÃO DE MICRODADOS DOS CENSOS EDUCACIONAIS

5.4. É importante esclarecer que o *Guia Orientativo sobre o Tratamento de Dados Pessoais pelo Setor Público*, publicado pela ANPD, em janeiro de 2022, no sítio eletrônico <https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/guia-poder-publico-anpd-versao-final.pdf>, aponta que o processo de adequação às disposições da LGPD tem suscitado muitas dúvidas a respeito dos parâmetros a serem observados para a disponibilização pública de informações pessoais, sobretudo quanto à ponderação entre direitos: de um lado, o direito à privacidade e o direito à proteção de dados pessoais e, de outro, o direito coletivo à informação sobre as atividades do Poder Público.

5.5. Como é de conhecimento, a Lei nº 12.527, de 17/11/2011 - Lei de Acesso à Informação (LAI), estabeleceu a determinação legal de que a publicidade da informação é a regra, admitindo-se o sigilo apenas em hipóteses excepcionais. Embora a observância da publicidade seja preceito geral para as informações detidas por órgãos e entidades do poder público, quando se trata de dados pessoais, faz-se necessário que o tratamento desses dados pessoais pelo Poder Público, incluindo a divulgação pública de dados pessoais, seja realizado em conformidade com os preceitos da LGPD, sobretudo os dispositivos que garantem a proteção integral dos dados pessoais, a autodeterminação informativa e o respeito à privacidade dos titulares durante todo o ciclo do tratamento. Dessa forma, desde a realização da coleta até o fim da atividade, entidades e órgãos públicos devem observar os princípios previstos na lei, verificar a existência de base legal aplicável ao tratamento, assegurar os direitos dos titulares e adotar medidas de prevenção e segurança, a fim de evitar a ocorrência de incidentes.

5.6. Não se pode negar que há interesse público na divulgação de informações relativas à execução de políticas públicas e ao desempenho de competências legais pelos órgãos e entes públicos que permitam aos cidadãos o exercício do controle social. Sendo assim, com vistas a atender ao princípio da publicidade, com frequência, o Estado é obrigado a divulgar dados pessoais.

5.7. Nesse contexto, o cumprimento da LGPD demanda de entidades e órgãos públicos uma análise mais ampla, que não se limita à atribuição de sigilo ou de publicidade a determinados dados pessoais. Faz-se mister, em vista do reforço protetivo trazido pela LGPD, uma avaliação sobre os riscos e os impactos para os titulares dos dados pessoais, bem como sobre as medidas mais adequadas para mitigar possíveis danos decorrentes do tratamento.

5.8. Feitas essas considerações, cumpre aplicá-las ao caso sob análise: divulgação de microdados dos censos escolares e do Enem no sítio eletrônico do INEP.

5.9. O Decreto nº 6.425/2008^[2] determina que o INEP realizará, anualmente, o censo escolar da educação básica e o censo da educação superior. O censo é realizado em regime de colaboração entre a União, os Estados, o Distrito Federal e os Municípios, em caráter declaratório e mediante coleta de dados descentralizada, englobando todos os estabelecimentos públicos e privados de educação básica e adotando alunos, turmas, escolas e profissionais da educação como unidades de informação. O art. 5º do referido decreto impõe a obrigatoriedade a toda instituição de educação, de direito público ou privado, com ou sem fins lucrativos, de prestar as informações solicitadas pelo INEP por ocasião da realização do censo da educação ou para fins de elaboração de indicadores educacionais. **Já o art. 6º da norma assegura o sigilo e a proteção de dados pessoais no censo da educação, vedando a utilização dos dados para fins estranhos aos previstos na legislação educacional aplicável.**

Decreto nº 6.425/2008 - Dispõe sobre o censo anual da educação.

Art. 5º Toda instituição de educação, de direito público ou privado, com ou sem fins lucrativos, é obrigada a prestar as informações solicitadas

pelo INEP, por ocasião da realização do censo da educação ou para fins de elaboração de indicadores educacionais.
Art. 6º Ficam assegurados o sigilo e a proteção de dados pessoais apurados no censo da educação, vedada a sua utilização para fins estranhos aos previstos na legislação educacional aplicável.

5.10. Quanto ao Enem, importa registrar que a Portaria MEC nº 807, de 18/06/2010^[3], instituiu o procedimento de avaliação com o objetivo de aferir se o participante, ao final do ensino médio, demonstra domínio dos princípios científicos e tecnológicos que presidem a produção moderna e conhecimento das formas contemporâneas de linguagem, a teor do art. 1º. No art. 6º da Portaria, por sua vez, está prevista a estruturação de banco de dados e emissão de relatórios com os resultados do Enem, bem como a possibilidade de disponibilização dos resultados do exame a instituições de ensino superior, secretarias estaduais de educação e pesquisadores, resguardado o sigilo individual, condicionando a divulgação do resultado individual à autorização expressa do participante. A restrição existe, portanto, à divulgação do resultado individual, para a qual deve existir autorização expressa do titular.

Portaria MEC nº 807, de 18/06/2010

Art. 6º O INEP estruturará um banco de dados e emitirá relatórios com os resultados do ENEM.

§ 1º A participação no ENEM conferirá ao examinando um boletim contendo informações referentes ao resultado global e ao resultado individual.

§ 2º Os resultados do ENEM, resguardado o sigilo individual, estarão disponíveis para instituições de ensino superior, secretarias estaduais de educação e pesquisadores.

§ 3º Os resultados individuais do ENEM somente poderão ser divulgados mediante a autorização expressa do participante

5.11. O INEP está enquadrado como entidade do Poder Público, motivo pelo qual são-lhe aplicáveis as orientações e os entendimentos estabelecidos no *Guia Orientativo sobre o Tratamento de Dados Pessoais pelo Setor Público* mencionado anteriormente. Verifica-se, também, que existe respaldo em lei^[4] para a divulgação de dados referentes à educação básica e superior.

5.12. Importante considerar, ainda, o disposto na LAI, cujo art. 31 prevê que o tratamento das informações pessoais seja feito de forma transparente e com respeito à intimidade, vida privada, honra e imagem das pessoas, bem como às liberdades e garantias individuais. Chama-se a atenção, por oportuno, que a Emenda Constitucional nº 115/2022, incluiu o direito à proteção de dados pessoais como uma garantia fundamental, em acréscimo à proteção constitucional já existente à intimidade e à vida privada. Destaque-se, ainda, o §1º, II do mesmo artigo, que estabelece que as informações pessoais poderão ter sua divulgação ou acesso por terceiros autorizado diante de previsão legal ou consentimento expresso da pessoa a que elas se referirem.

Lei nº 12.527/2011 - LAI

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.

Constituição Federal de 1988

Art. 5º Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade, nos termos seguintes:

(...)

X - são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação;

(...)

LXXIX - é assegurado, nos termos da lei, o direito à proteção dos dados pessoais, inclusive nos meios digitais.

5.13. Por sua vez, o art. 31, §3º, II e V, da LAI afastam a necessidade de consentimento para a divulgação de informações pessoais quando necessárias "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"; e "à proteção do interesse público e geral preponderante". Tendo em vista que os microdados do censo escolar e do Enem se afiguram como necessários à realização de estatísticas e pesquisas científicas, conforme explanado anteriormente, e também refletem proteção do interesse público geral e preponderante, pois auxiliam na execução de política pública de avaliação do sistema brasileiro de ensino; a situação, a princípio, atende aos requisitos previstos na LAI para a divulgação pública sem a necessidade de consentimento dos titulares.

Lei nº 12.527/2011 - LAI

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.

(...)

§ 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.

5.14. Com base no exposto, verifica-se que o tratamento de dados pessoais pelo INEP para fins de censo escolar e de divulgação do Enem está respaldado em lei, pois é desenvolvido no mister do cumprimento de obrigações legais atribuídas à Autarquia ou, ainda, para fins da execução da política pública de acompanhamento da educação básica e superior e de avaliação do sistema de ensino brasileiro. Dito isso, do ponto de vista da LGPD, e com base nas notas técnicas do INEP, no Parecer 00018/2022/PROC/PFINEP/PGF/AGU (SEI nº 3289237) e no entendimento já manifestado pela CGF em sua Nota Técnica 46 (SEI nº 3319546), é possível concluir que a hipótese legal mais adequada para o tratamento de dados pessoais pelo INEP não deve ser a do consentimento. Isso porque a referida lei autoriza de forma expressa o tratamento de dados com a finalidade de realização de estudos por órgão de pesquisa, ou ainda para o cumprimento de obrigação legal ou regulatória, conforme disposto nos artigos 7º e 11 da LGPD. Não restam dúvidas, portanto, acerca da legitimidade do tratamento de dados realizado pelo INEP para fins de realização dos censos educacionais e divulgação das informações pertinentes aos interessados.

III. DIVULGAÇÃO PÚBLICA DE DADOS PESSOAIS E ANONIMIZAÇÃO

5.15. Inicialmente, vale destacar que a LGPD, ao reconhecer a realização de estudos por órgão de pesquisa como uma das hipóteses legais que legitima o tratamento de dados pessoais sem a necessidade de consentimento do titular, também aduziu, no mesmo dispositivo e de forma objetiva, a garantia da anonimização dos dados pessoais sensíveis, sempre que possível. De forma análoga, a própria LAI, em seu art. 31, oferece salvaguardas para a tensão entre o interesse público - materializado na necessidade de produção das estatísticas oficiais-, e o interesse individual do cidadão de não ter sua intimidade revelada. Ao passo que dispensa o consentimento do titular de dados para a realização de estatísticas e pesquisas

científicas de evidente interesse público ou geral, previstos em lei, a LAI veda, no mesmo dispositivo, a identificação da pessoa a que as informações se referirem.

5.16. Dado pessoal anonimizado, conforme disposto no art. 5º, III, da LGPD, é o dado relativo a um titular que não possa ser identificado. Em outras palavras, é o dado pessoal tornado anônimo de forma que não seja possível a identificação do seu titular, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento. Ou seja, a anonimização é o processo por meio do qual um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo, tornando-se, portanto, anonimizado, conforme definido no art. 5º, XI, da LGPD. □

5.17. A reidentificação, por outro lado, é a possibilidade de identificar um indivíduo único, transformando dados anônimos em dados pessoais por meio do uso de correspondência de dados ou técnicas semelhantes. O processo de anonimização e a forma como é implementado terão uma influência direta na probabilidade de ocorrer reidentificação. Considerando que, para se anonimizar um dado pessoal, serão utilizados meios técnicos razoáveis e disponíveis no momento desse processo, existe o risco de que alguns processos de anonimização possam ser revertidos no futuro. Além disso, é preciso reconhecer que as circunstâncias que favorecem a reidentificação podem mudar com o tempo e que novos desenvolvimentos tecnológicos e a disponibilidade de informações adicionais podem comprometer os processos de anonimização anteriores.

5.18. Também é importante registrar que a anonimização não reduz a probabilidade de reidentificação de um conjunto de dados a zero. Embora a anonimização total seja o objetivo desejável do ponto de vista da proteção de dados pessoais, em alguns casos, isso não é possível. Assim, deve ser considerado um risco residual de reidentificação. No âmbito da Administração Pública, a avaliação de riscos e de impactos aos titulares dos dados, bem como a adoção de salvaguardas e de medidas de mitigação proporcionais aos riscos identificados devem ser um dos condicionantes para a decisão do controlador na publicização de dados pessoais.

5.19. Não se pode negar que existem riscos e interesses em reidentificar os dados anonimizados. Os dados pessoais têm um valor em si, para os próprios indivíduos e para terceiros. A reidentificação de um indivíduo pode ter um sério impacto sobre seus direitos e liberdades. Assim, a possibilidade de se reidentificar uma pessoa em um conjunto de dados, seja por curiosidade, por acaso ou motivado por um interesse legítimo (como, por exemplo, para pesquisa científica e para fins jornalísticos) ou ilegítimo (como, por exemplo, para atividade criminosa), não pode ser desconsiderada.

5.20. O conjunto de dados que inclui dados pessoais pode conter identificadores que possibilitam a associação, direta ou indireta, a um indivíduo, nos termos do art. 5º, XI e art. 12, § 1º, da LGPD. Um identificador direto é uma informação específica que se refere a um indivíduo, como nome ou número de registro. Um identificador indireto é qualquer informação que pode ser usada, individualmente ou em combinação com outras informações, com o objetivo de identificar uma pessoa no conjunto de dados, como, por exemplo, a sua localização geográfica, um endereço IP, uma opinião em redes sociais, ou ainda aquelas utilizadas para a formação do perfil comportamental de uma pessoa, se identificada, conforme o art. 12, § 2º, da LGPD.

5.21. Como parte da metodologia de anonimização, utilizam-se medidas para garantir que o limite de risco de reidentificação não seja ultrapassado, como por exemplo o k-anonimato. O modelo k-anonimato é usado como uma diretriz antes e após as técnicas de anonimização terem sido aplicadas, para garantir que os identificadores diretos e/ou indiretos (os "quase identificadores") de qualquer registro sejam compartilhados por pelo menos k-1 outros registros. "K" é a proteção de chave fornecida contra ataques de vinculação, onde k registros são idênticos nos atributos de identificação e, portanto, criam uma "classe de equivalência" com k membros, não sendo possível vincular ou isolar o registro de um único indivíduo sempre que há "k" indivíduos com atributos idênticos.

5.22. Um conjunto de dados anonimizados pode ter diferentes níveis de k-anonimato para diferentes conjuntos de identificadores indiretos (quase identificadores), mas para obter a proteção contra vinculação, o k não poderá ser um número baixo que comprometa a anonimização. Pelo exposto, considerando o caso em análise, à luz das informações trazidas no Termo de Execução Descentralizada UFMG (SEI nº 3289230), observa-se que não era necessário possuir acesso a outras fontes de dados abertos (não anonimizados) para se obter a reidentificação de registros anonimizados quando publicados no formato anterior, especialmente quando o fator de anonimização for baixo, o que sugere uma maior facilidade de reidentificação dos titulares.

5.23. Importante frisar que os dados anonimizados não são considerados dados pessoais, por isso não estão sujeitos às disposições da LGPD, salvo quando o processo de anonimização a que foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido (art. 12). A determinação do que seja esforço razoável deve levar em consideração fatores objetivos, tais como o custo e o tempo necessários para reverter o processo de anonimização, de acordo com as tecnologias disponíveis no momento da anonimização, e a utilização exclusiva de recursos tecnológicos próprios do agente de tratamento (art. 12, § 1º).

5.24. Ressalte-se que a anonimização não é uma medida de segurança impositiva, que deve ser adotada em todo e qualquer tratamento de dados pessoais. De forma mais específica, a LGPD não estabeleceu a anonimização como condição técnica para a divulgação pública ou para o compartilhamento de dados pessoais por entidades e órgãos públicos. Como já mencionado, a principal determinação da LGPD é quanto à necessidade de avaliação de riscos e de adoção de medidas para mitigar a ocorrência de danos. Por isso, a eventual identificação dos titulares ou a admissão de algum grau de risco de sua identificação, quando necessário para atender, por exemplo, a determinações legais, o interesse público e o direito de acesso à informação, são compatíveis com a LGPD, desde que adotadas as salvaguardas apropriadas.

5.25. Deve-se registrar que a preocupação com a anonimização no presente caso não advém da LGPD, mas sim dos mecanismos adotados pelo próprio INEP em suas pesquisas. Isso porque a lógica do modelo de pesquisa desenvolvido exige que ele seja impessoal, conforme ressaltado nas Notas Técnicas do INEP, no bojo das quais se afirma que "a confidencialidade dos dados pessoais tratados é um pressuposto ético e um requisito metodológico". Não se trata, portanto, de uma obrigação que surgiu com o advento da LGPD.

5.26. De fato, a anonimização é uma das possíveis medidas de segurança que podem ser adotadas visando à proteção de dados pessoais. Ainda no âmbito das medidas técnicas, pode ser mencionada a pseudonimização, que é definida pela LGPD como "o tratamento por meio do qual um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo, senão pelo uso de informação adicional mantida separadamente pelo controlador em ambiente controlado e seguro" (art. 13, § 4º). Embora a única menção ao termo seja efetuada no art. 13, que trata dos estudos em saúde pública, a pseudonimização de dados pessoais, como, por exemplo, por meio de criptografia, pode ser utilizada nos mais diversos contextos, sempre considerando os objetivos de prevenção e segurança identificados no caso concreto.

5.27. Além dos meios técnicos, a segurança no tratamento de dados pessoais pode ser garantida mediante a adoção de ações administrativas, a exemplo da definição de critérios mais restritos para controle e registro de acesso a determinados dados pessoais, conforme previsto na política de segurança da informação do controlador. No âmbito da divulgação de dados pessoais por entidades e órgãos públicos, a eventual limitação do acesso a pessoas previamente cadastradas ou que se enquadrem em determinadas categorias (como pesquisadores, gestores públicos etc.) pode ser útil em determinados contextos, observadas as disposições legais aplicáveis. Da mesma forma, mecanismos jurídicos, como a assinatura de termos de responsabilidade pelas pessoas que têm acesso aos dados pessoais, podem contribuir para a mitigação de riscos aos titulares dos dados. Nesse sentido, vale citar o disposto no art. 61 do Decreto nº 7.724/2012, que regulamenta a aplicação da LAI no Poder Executivo Federal: □

Decreto nº 7.724/2012

Art. 61. O acesso à informação pessoal por terceiros será condicionado à assinatura de um termo de responsabilidade, que disporá sobre a

finalidade e a destinação que fundamentaram sua autorização, sobre as obrigações a que se submeterá o requerente.

§ 1º A utilização de informação pessoal por terceiros vincula-se à finalidade e à destinação que fundamentaram a autorização do acesso, vedada sua utilização de maneira diversa.

§ 2º Aquele que obtiver acesso às informações pessoais de terceiros será responsabilizado por seu uso indevido, na forma da lei.

5.28. Como se pode observar, o termo de responsabilidade é instrumento jurídico utilizado para dar ciência expressa ao terceiro de sua corresponsabilidade pelo tratamento dos dados disponibilizados pelo controlador. Trata-se, em outras palavras, de estabelecer constrangimentos jurídicos, que se somam aos meios técnicos adotados no caso, visando assegurar que os dados pessoais sejam tratados em conformidade com as condições estipuladas para o acesso aos dados pessoais, entre as quais podem ser mencionadas: a boa-fé, o interesse público, a finalidade específica e o respeito aos direitos dos titulares e aos padrões éticos aplicáveis.

5.29. A existência de diversas salvaguardas técnicas e administrativas, admitidas e compatíveis com a LGPD e que não se limitam à técnica de anonimização, conforme exemplificado acima, impõe aos controladores em geral, e ao INEP, em particular, a obrigação de realizar uma avaliação ampla dos riscos envolvidos e das medidas de segurança e de mitigação de riscos mais apropriadas para o caso.

5.30. Para tanto, o instrumento previsto na LGPD é o Relatório de Impacto à Proteção de Dados Pessoais (RIPD), definido no art. 5º, XVII, como a “documentação do controlador que contém a descrição dos processos de tratamento de dados pessoais que podem gerar riscos às liberdades civis e aos direitos fundamentais, bem como medidas, salvaguardas e mecanismos de mitigação de risco”. O parágrafo único do art. 38 estabelece o conteúdo mínimo do relatório, que deve conter, pelo menos, “a descrição dos tipos de dados coletados, a metodologia utilizada para a coleta e para a garantia da segurança das informações e a análise do controlador com relação a medidas, salvaguardas e mecanismos de mitigação de risco adotados”.

IV. RELATÓRIO DE IMPACTO À PROTEÇÃO DE DADOS PESSOAIS

5.31. Como mencionado, a análise do controlador em relação a “medidas, salvaguardas e mecanismos de mitigação de risco” não precisa se limitar à técnica de anonimização, devendo também considerar a pertinência da aplicação de outras iniciativas técnicas e administrativas e a devida ponderação entre o interesse público na divulgação dos dados e a proteção dos direitos dos titulares. No presente caso, a elaboração do relatório de impacto se justifica: (I) em razão do risco já identificado aos titulares em decorrência da publicação das bases de dados dos censos educacionais; e (II) da necessidade de reforçar a conformidade do tratamento dos dados realizado por INEP com a LGPD.

5.32. A esse respeito, o art. 32 da LGPD expressamente atribui à ANPD a competência para “solicitar a agentes do Poder Público a publicação de relatórios de impacto à proteção de dados pessoais e sugerir a adoção de padrões e de boas práticas para os tratamentos de dados pessoais pelo Poder Público”. Dessa forma, no âmbito do setor público, o relatório de impacto ganha contornos próprios, na medida em que: (I) viabiliza uma atuação mais efetiva da ANPD no caso concreto, pois considerando a análise e as informações apresentadas pelo INEP, poderá recomendar ou determinar a adoção de providências adicionais; e (II) contribui para ampliar a transparência dos critérios adotados e considerados pelo INEP para fundamentar a sua decisão, em particular mediante a eventual publicação do relatório, caso assim determine a ANPD.

5.33. Conforme consignado na Nota Técnica 46 (SEI nº 3319546) da CGF, a existência de respaldo legal para a divulgação dos microdados não significa que ela deva continuar a ser feita nos moldes antigos. É preciso observar os princípios previstos na LGPD, sobretudo o da necessidade, de sorte a publicar apenas os dados pertinentes ao escopo de avaliação do ensino brasileiro e da realização de pesquisas estatísticas educacionais. Com efeito, o INEP precisa atender à lógica da necessidade, pois a autorização legal para publicar os dados não redundaria numa divulgação que ignore a necessidade e a adequação. Isto porque o advento da LGPD impõe a observância de princípios e dos direitos do titular, promovendo uma cultura de melhoria no tratamento de dados por meio da positivação, no inciso I do art. 6º, do princípio da finalidade, de modo que o tratamento esteja centrado no mínimo imprescindível; do princípio da adequação, no inciso II, que impõe a compatibilidade do tratamento com as finalidades informadas ao titular, de acordo com o contexto do tratamento; do princípio da necessidade, no inciso III, para que o tratamento limite-se ao mínimo necessário para a realização de suas finalidades, com abrangência dos dados pertinentes, proporcionais e não excessivos em relação às finalidades. Assim, a finalidade determina o que se pode fazer com os dados.

5.34. Em resumo, o INEP, que está amparado em hipótese legal para o tratamento dos microdados, deve atender aos princípios da finalidade, adequação e necessidade, verificando se as informações divulgadas são, efetivamente, adequadas e necessárias para o atendimento da finalidade a que o Instituto se propõe com a divulgação (tratamento) dos dados.

5.35. Nesse ponto, importante ressaltar que a própria divulgação consiste em tipo de tratamento previsto no art. 5º, X, da LGPD: “toda operação realizada com dados pessoais, como as que se referem a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração”. Assim, já que a divulgação configura tratamento, faz-se mister que ela seja feita com respeito aos princípios previstos pela LGPD, de sorte a garantir que a finalidade do tratamento seja atendida. Nessa linha de raciocínio, o fato de os dados tornados públicos poderem ser usados sem consentimento (desde que observados os §§ 3º, 4º e 7º do art. 7º da LGPD) deve inspirar a reflexão sobre a publicação, a fim de que sejam divulgados apenas os dados efetivamente necessários e compatíveis com a realização de estatísticas e pesquisas científicas e a execução da política pública de avaliação do sistema brasileiro de ensino que gerou o tratamento.

5.36. Miriam Wimmer, Diretora da ANPD, no artigo “O regime jurídico do tratamento de dados pessoais pelo Poder Público”^[5], sugere a realização de relatórios de impacto à proteção de dados pessoais antes da abertura dos dados a serem publicados pelo Poder Público, e este é o posicionamento adotado pela ANPD no presente caso. Nessa senda, devem ser levados em consideração quais impactos a divulgação dos microdados pode causar aos direitos fundamentais ou no que a exposição pode afetar a vida dos indivíduos. No caso de dados identificáveis que podem, efetivamente, sujeitar o titular a discriminação, por exemplo, a publicação deve ser evitada e, em sendo necessária para política pública, o acesso deve ocorrer em ambiente seguro ou mediante a adoção de outras salvaguardas apropriadas. Assim, se o RIPD concluir que os dados são pessoais porque não estão anonimizados, entre a deliberação pela publicação e a publicação em si, deve haver reflexão sobre o que se precisa fazer. O documento, então, informará as precauções que a autarquia federal tomará no intuito de mitigar os riscos às liberdades civis e aos direitos fundamentais.

5.37. Maria Cecília Oliveira Gomes, sobre a descrição dos processos de tratamento de dados no relatório de impacto, ensina que “parece que a intenção do legislador não foi que o relatório fosse uma demonstração do registro de atividades ou um inventário de dados especificamente, o qual também está previsto na LGPD e no Marco Civil da Internet, mas sim que contivesse uma descrição dos procedimentos que envolvem as operações de tratamento de dados, através de uma perspectiva maior de governança de dados”^[6]. Para além da governança de dados no presente, acrescente-se a necessidade de se adotar uma postura prospectiva, pois se for cogitada a divulgação de novos tipos de dados, deve ser elaborado novo RIPD. Agindo dessa maneira, o INEP promoverá continuamente a adequação do seu tratamento à LGPD, refletindo a postura de cautela que é almejada pela nova legislação.

5.38. Por fim, com a elaboração do RIPD, atende-se também ao princípio da prevenção, disposto no art. 6º, VIII, segundo o qual devem ser adotadas medidas para prevenir a ocorrência de danos em virtude do tratamento de dados pessoais. A elaboração do RIPD permitirá, portanto, uma melhor avaliação sobre a forma como os microdados tratados pelo INEP poderão ser, ou não, tornados públicos.

IV.1 RIPD dos microdados do Exame Nacional do Ensino Médio (Enem)

5.39. Em atenção aos encaminhamentos propostos pela CGF em sua Nota Técnica 46 (SEI nº 3319546), o INEP apresentou o Relatório de Impacto de Proteção de dados do Exame Nacional do Ensino Médio, ENEM (Relatório RIPD Microdados do Enem, SEI nº 3848205). É esse documento que se passa a analisar no presente tópico.

5.40. Conforme publicado pela ANPD em seu sítio eletrônico^[7], é recomendável que o RIPD reúna os seguintes dados e informações:

- a) Identificação dos agentes de tratamento e do encarregado;
- b) Outras partes interessadas/envolvidas. Informar se foram consultadas na elaboração do RIPD e pareceres emitidos;
- c) Justificativa da necessidade de elaboração do relatório (por exemplo: alto risco, solicitação da ANPD, gestão de riscos e prevenção, outros);
- d) Projeto/Processo que justifica a elaboração do RIPD;
- e) Sistemas de informação relacionados ao projeto/processo que justifica a elaboração do RIPD;
- f) Tratamento de dados;
 - i. Descrição do tratamento (desde a coleta até a eliminação);
 - ii. Dados pessoais (informar todos os tipos de dados pessoais tratados, de forma completa);
 - iii. Dados pessoais sensíveis (informar todos os tipos de dados pessoais sensíveis tratados, de forma completa);
 - iv. Categorias de titulares (por exemplo, clientes, funcionários do controlador, filhos de funcionários do controlador, funcionários de clientes, autores de ações judiciais, beneficiários de apólices, terceiros prestadores de serviços);
 - v. Dados de crianças e adolescentes ou de outra categoria de vulneráveis, como idosos, se houver;
 - vi. Volume de dados pessoais tratados e número de titulares envolvidos no tratamento;
 - vii. Fonte de coleta;
 - viii. Finalidade do tratamento (Justifique a finalidade de tratamento para cada dado);
 - ix. Informar quais são os compartilhamentos internos e externos (inclusive transferência internacional, se houver);
 - x. Política de armazenamento (descrever os prazos de retenção e métodos de descarte);
- g) Análise de hipótese legal. Justifique a escolha da hipótese legal para cada finalidade de tratamento;
- h) Análise de princípios da LGPD;
- i) Riscos identificados ao titular;
- j) Resultado apurado com base na metodologia utilizada pelo agente de tratamento.

5.41. Em relação ao primeiro tópico (a), observa-se que o RIPD indicou o próprio INEP como controlador e operador de dados, sendo que o papel de operador é designado à Diretoria de Avaliação da Educação Básica (DAEB). Desse modo, entende-se que o INEP não somente toma as decisões referentes ao tratamento de dados pessoais (controlador), mas também é aquele que realiza o tratamento de dados pessoais (operador)^[8], por meio de uma diretoria específica. Verifica-se, também, que o RIPD indicou corretamente o encarregado, sua identificação e informações de contato (e-mail e telefone).

5.42. No tocante ao item (b) outras partes interessadas/envolvidas, o INEP mencionou o envolvimento da própria ANPD, por meio de reunião e também nos autos do presente processo, consubstanciado nas orientações e encaminhamentos propostos na Nota Técnica 46 (SEI nº 3319546). Também foram endereçadas as preocupações levantadas pela CGU, em sua Nota Técnica, acerca de aspectos gerais envolvidos na publicação dos microdados.

5.43. Como justificativa da necessidade (c), o relatório aponta a solicitação da ANPD por meio da Nota Técnica 46 (SEI nº 3319546), enquadrando-se na hipótese do art. 38 da LGPD. Muito embora tal justificativa seja válida e suficiente, importa registrar, a título de esclarecimento, que cabe ao controlador gerir os riscos associados a todo tratamento de dados pessoais que realiza. Considerando-se que o RIPD abriga "a análise do controlador com relação a medidas, salvaguardas e mecanismos de mitigação de risco adotados", importa chamar atenção para o constante na página da ANPD sobre quando é recomendável pela ANPD a elaboração do RIPD^[9]:

3. Em qual contexto a ANPD recomenda que seja elaborado o RIPD?

Como regra geral, é recomendado elaborar o RIPD em todo contexto em que as operações de tratamento de dados pessoais possam gerar alto risco à garantia dos princípios gerais de proteção de dados pessoais previstos na LGPD e às liberdades civis e aos direitos fundamentais do titular de dados, conforme art. 5º, inciso XVII, e art. 55-J, inciso XIII, da LGPD, o que deverá ser avaliado pelo agente de tratamento.

A LGPD lista, ainda, situações específicas em que o RIPD poderá ser exigido pela ANPD, como:

- nas operações de tratamento efetuadas para fins exclusivos de segurança pública, defesa nacional, segurança do Estado ou atividades de investigação e repressão de infrações penais (art. 4º, § 3º);
- quando o tratamento tiver como fundamento a hipótese de interesse legítimo (art. 10, § 3º);
- para agentes do Poder Público, incluindo determinação quanto à publicação do RIPD (art. 32); e
- para controladores em geral, quanto às suas operações de tratamento, incluindo as que envolvam dados pessoais sensíveis (art. 38).

Portanto, **haverá situações em que o controlador elaborará o RIPD para atender à determinação da ANPD ou, em atenção ao princípio da responsabilização e prestação de contas (art. 6º, X), ao verificar que o tratamento a ser realizado pode implicar alto risco à garantia dos princípios gerais de proteção de dados pessoais previstos na LGPD e às liberdades civis e aos direitos fundamentais do titular de dados.**

Além disso, a LGPD prevê a possibilidade de que os controladores, para cumprimento dos princípios da segurança e da prevenção (art. 6º, VII e VIII), implementem programa de governança em privacidade que, entre outros itens, estabeleça políticas e salvaguardas adequadas com base em processo de avaliação sistemática de impactos e riscos à privacidade (art. 50, § 2º, I, d), procedimento que pode envolver a elaboração de RIPD.

5.44. Em atenção ao item (d), o INEP indica, como processo que justifica a confecção do RIPD, a elaboração e divulgação dos microdados do ENEM, que envolve o tratamento de dados pessoais dos participantes de cada edição desse exame educacional.

5.45. Como sistema de informação relacionado ao processo de elaboração e divulgação dos microdados do ENEM (e), o RIPD indica o software estatístico SAS Analytics Software.

5.46. Na sequência, o relatório descreve como é realizado o tratamento dos dados pessoais coletados dos participantes do exame (item f). De acordo com o INEP, os dados coletados são agrupados em tabelas, conhecidas internamente no INEP como Tabelas Simplificadas (TS), e

compõem uma base de dados denominada DATA_ENEM. Essas tabelas consistem na principal fonte de dados do processo de geração dos microdados. Os demais itens que compõem o pacote de microdados são construídos (dicionário de dados, tabela de itens com os parâmetros, programas de INPUT e demais documentos). O produto informacional é então compactado em arquivo único e disponibilizado para download público no portal do INEP.

5.47. O INEP informa que a quantidade de dados pessoais tratados é de 76, sendo 1 dado pessoal sensível (cor/raça). A frequência de realização desse processo de tratamento dos dados pessoais é anual. Após a divulgação dos resultados do ENEM, inicia-se o processo de elaboração dos microdados, com o objetivo de divulgá-los até o final de junho de cada ano. O número de titulares afetados pelo tratamento varia conforme a edição do exame. Em 2021, última edição em que os microdados foram divulgados, havia na base de dados registros de 3.389.832 participantes do ENEM, atingindo volume de 1,4 GB de dados pessoais.

5.48. O RIPD informa que os arquivos que compõem o pacote do microdados do ENEM são armazenados permanentemente em estruturas de pastas na rede corporativa, enquanto se mantiver a divulgação dos microdados do exame. A estrutura de pastas na rede utilizada para armazenar esses arquivos é acessível apenas pelos servidores da coordenação responsável. As medidas de segurança atualmente adotadas são: Controle de Acesso Lógico, Controles de Segurança em Redes e Proteção Física e do Ambiente.

5.49. De acordo com o relatório apresentado, os dados pessoais tratados no processo de elaboração e divulgação dos microdados do ENEM abrangem:

- Informações do Participante: número da inscrição (máscara), ano de aplicação do ENEM, faixa etária, sexo, estado civil, cor/raça, nacionalidade, situação de conclusão do ensino médio, ano de conclusão do ensino médio, tipo de escola do ensino médio, tipo de instituição em que concluiu ou concluirá o ensino médio, indicação de inscrito treineiro.
- Informações da Escola: código do município da escola, nome do município da escola, código da UF da escola, sigla da UF da escola, dependência administrativa, localização, situação de funcionamento.
- Informações do Local de Aplicação das Provas: código do município da aplicação da prova, nome do município da aplicação da prova, código da UF da aplicação da prova, sigla da UF da aplicação da prova.
- Informações da Prova Objetiva: presença na prova objetiva de Ciências da Natureza, presença na prova objetiva de Ciências Humanas, presença na prova objetiva de Linguagens e Códigos, presença na prova objetiva de Matemática, código do tipo de prova de Ciências da Natureza, código do tipo de prova de Ciências Humanas, código do tipo de prova de Linguagens e Códigos, código do tipo de prova de Matemática, nota da prova de Ciências da Natureza, nota da prova de Ciências Humanas, nota da prova de Linguagens e Códigos, vetor com as respostas da parte objetiva da prova de Ciências da Natureza, vetor com as respostas da parte objetiva da prova de Ciências Humanas, vetor com as respostas da parte objetiva da prova de Linguagens e Códigos, vetor com as respostas da parte objetiva da prova de Matemática, opção de língua Estrangeira, vetor com o gabarito da parte objetiva da prova de Ciências da Natureza, vetor com o gabarito da parte objetiva da prova de Ciências Humanas, vetor com o gabarito da parte objetiva da prova de Linguagens e Códigos, vetor com o gabarito da parte objetiva da prova de Matemática.
- Informações da Redação: situação da redação do participante, nota da competência 1, nota da competência 2, nota da competência 3, nota de competência 4, nota de competência 5, nota da prova de redação.
- Informações do Questionário Socioeconômico: respostas a vinte e cinco questões do questionário socioeconômico, preenchido pelo participante no ato da inscrição.

5.50. O relatório esclarece que há um registro para cada participante na base, e que o número real da inscrição foi substituído por uma máscara. Acrescenta que não há, nos microdados disponibilizados publicamente, quaisquer informações que permitam a identificação direta do participante, tais como CPF, nome, nome da mãe, data de nascimento, logradouro do endereço, e-mail etc. Adicionalmente, o INEP informa que poderá ocorrer tratamento de dados pessoais de adolescentes, já que esse público-alvo poderá participar do ENEM na condição de treineiro.

5.51. No que tange à finalidade de tratamento, o relatório afirma que a base de microdados gerada a partir dos dados indicados acima é um dos produtos informacionais mais relevantes produzidos pelo Instituto, uma vez que reúne um conjunto de informações detalhadas utilizadas como subsídios a pesquisas e avaliações do Instituto. O INEP destaca o elevado interesse público dessa base, utilizada por pesquisadores, representantes de instituições privadas, gestores governamentais e outros interessados na área da educação para realizar análises, diagnósticos, estudos e pesquisas de interesse social.

5.52. Embora o RIPD não dedique um tópico para abordar a análise da hipótese legal para o tratamento dos microdados do ENEM (item g), o relatório faz referência à Nota Técnica 46 (SEI nº 3319546) da CGF, em que, nos termos do RIPD ora em análise, a ANPD "reconheceu que o INEP possui hipótese legal para realizar a divulgação dos microdados." Com efeito, da leitura da referida Nota Técnica da CGF, verifica-se que o tratamento de dados pessoais pelo INEP no âmbito da divulgação dos microdados do Enem está respaldado pela LGPD, pois é desenvolvido no mister do cumprimento de obrigações legais atribuídas à autarquia ou, ainda, para fins da execução da política pública e realização de estudos de acompanhamento do sistema de ensino brasileiro. O tratamento de dados pessoais pelo INEP para a elaboração e posterior divulgação dos microdados do ENEM está respaldado pelos inc. II e IV do art. 7º da LGPD. Não há dúvidas, portanto, quanto à legitimidade do tratamento de dados realizado pelo INEP no caso em análise.

5.53. Acerca da (h) análise de princípios da LGPD, o RIPD foca no aspecto da necessidade. O relatório reconhece que, segundo o princípio da necessidade, o tratamento deve ser limitado ao "mínimo necessário para a realização de suas finalidades", abrangendo apenas os "dados pertinentes, proporcionais e não excessivos em relação às finalidades do tratamento de dados" (art. 6º, III LGPD). O princípio da necessidade impõe, portanto, que a coleta (e qualquer outro tratamento) se atenha à menor quantidade possível de dados para o alcance da finalidade proposta.

5.54. Estudos solicitados pelo INEP demonstraram que algumas variáveis presentes nos microdados favoreciam a identificação indevida dos participantes dos exames e avaliações da educação básica, notadamente o ENEM. De acordo com o RIPD, "Há, inclusive, evidências concretas de que é possível identificar indevidamente os participantes, a partir dos microdados, utilizando filtros combinados por municípios de nascimento e residência, idade e escola de conclusão do ensino médio. Essas são informações que a maior parte das pessoas mantém em seus perfis públicos nas redes sociais, evidenciando um risco que não poderia ser desconsiderado."

5.55. A partir da constatação de risco de identificação indevida do titular, o INEP investigou quais variáveis favoreciam essa identificação, bem como analisou o princípio da necessidade aplicado a cada uma dessas variáveis coletadas. Com base nos resultados alcançados, o INEP propôs uma versão simplificada da base de microdados para divulgação pública. O modelo simplificado representou as seguintes alterações nos microdados do ENEM:

5.55.1. **Exclusão da variável "CO_ESCOLA":** essa variável corresponde ao código da escola onde o participante afirma ter concluído o ensino médio. O seu preenchimento é autodeclarado pelo participante e não é validada junto ao Censo Escolar. Para além do fato desta variável

contribuir para a indevida identificação do titular do dado, o INEP alega que essa variável vinha sendo utilizada pelo público para gerar *rankings* entre escolas de forma falha e pouco acurada, o que vai de encontro aos objetivos do exame e gerava diversas reclamações junto ao instituto.

5.55.2. **Exclusão de informações referentes aos pedidos de atendimento especializado e específico, e recursos de atendimento especializado e específico para a realização da prova:** o RIPD acertadamente classifica essas informações como dados sensíveis dos participantes. De acordo com o INEP, como o número de participantes que solicitam atendimento e recursos especializados ou específicos é muito pequeno se comparado com o número total de inscritos, essas variáveis facilitarão a identificação do participante. Dessa forma, o instituto considerou mais prudente remover tais variáveis dos microdados. Embora essa exclusão possa gerar críticas, o INEP ressalta que há nas sinopses estatísticas do ENEM resultados agregados por UF e região dos participantes que solicitaram atendimento especializado e específico e recursos de atendimento especializado e específico para a realização da prova. Dessa forma, viabilizam-se as pesquisas relacionadas a esse público, sem o risco de identificação indevida.

5.55.3. **Substituição de "Idade" por "Faixa Etária":** para dificultar a identificação do participante, a variável NU_IDADE foi substituída na base dos microdados por uma faixa etária, cujas categorias são as mesmas empregadas na sinopse estatística do ENEM (para fins de padronização). De acordo com o INEP, a experiência demonstra que a substituição da variável por uma faixa etária não prejudicará as pesquisas.

5.55.4. **Exclusão de informações referentes aos municípios de nascimento e residência do participante:** as variáveis CO_MUNICIPIO_RESIDENCIA, NO_MUNICIPIO_RESIDENCIA, CO_UF_RESIDENCIA e SG_UF_RESIDENCIA disponibilizam informações referentes ao município onde o participante declara residir. De maneira similar, as variáveis CO_MUNICIPIO_NASCIMENTO, NO_MUNICIPIO_NASCIMENTO, CO_UF_NASCIMENTO e SG_UF_NASCIMENTO disponibilizam informações referentes ao município onde o participante declara ter nascido. De acordo com os estudos do INEP, essas informações, combinadas com outros dados cadastrais eventualmente publicizados, facilitam a identificação indevida do participante. De toda forma, o INEP relata que manterá na base de microdados a informação sobre o município do local de prova do participante que, segundo o Instituto, é o dado mais relevante para os pesquisadores.

5.56. A experiência do INEP demonstra que as alterações realizadas para adaptação ao modelo simplificado resultaram na exclusão e/ou modificação de variáveis pouco relevantes para a realização de estudos e pesquisas educacionais. Segundo o RIPD, essas modificações reduziram significativamente o risco de identificação indevida dos participantes do exame, sem comprometer o princípio de transparência da informação e o acesso aos dados relevantes para a realização de estudos e pesquisas no setor.

5.57. Sobre o quesito i) riscos identificados ao titular, observa-se que o principal risco tratado pelo RIPD consiste na possibilidade de reidentificação do titular, embora haja outros. O quadro abaixo, extraído do RIPD, apresentam os riscos identificados e o respectivo nível de risco ponderado:

Quadro 1 - Risco referente ao tratamento de dados pessoais nos microdados do ENEM

Id	Risco referente ao tratamento de dados pessoais	P ¹	I ²	Nível de Risco (P x I) ³
R01	Modificação não autorizada	10	15	150
R02	Perda	5	15	75
R03	Falha/erro de processamento	5	15	75
R04	Reidentificação de dados pseudonimizados	15	10	150

Legenda: P – Probabilidade; I – Impacto.

¹ Probabilidade: chance de algo acontecer, não importando se definida, medida ou determinada objetiva ou subjetivamente, qualitativa ou quantitativamente, ou se descrita utilizando-se termos gerais ou matemáticos (ISO/IEC 31000:2009, item 2.19).

² Impacto: resultado de um evento que afeta os objetivos (ISO/IEC 31000:2009, item 2.18).

³ Nível de Risco: magnitude de um risco ou combinação de riscos, expressa em termos da combinação das consequências e de suas probabilidades (ISO/IEC 31000:2009, item 2.23 e IN SGD/ME nº 1, de 2019, art. 2º, inciso XIII).

Fonte: RIPD Relatório de Impacto à Proteção de Dados Pessoais dos Microdados do Exame Nacional do Ensino Médio (ENEM) (SEI 3848205)

5.58. Nota-se que, além do risco de identificação do titular, o risco de modificação de dados não autorizada também foi classificado como elevado.

5.59. O quadro abaixo apresenta as medidas propostas para tratar os riscos apontados, bem como o risco residual após a aplicação de tais medidas.

Quadro 2 - Medidas para tratar os riscos

Risco	Medida(s)	Efeito sobre o Risco ¹	Risco Residual ²			Medida(s) ³ Aprovada(s)
			P	I	Nível (P x I)	
R01 – Modificação não autorizada	Controle de acesso lógico, desenvolvimento seguro e segurança em redes	Reduzir	5	10	50	Sim
R02 – Perda	Backup do pacote de microdados, restrição de alteração das tabelas simplificadas apenas a administradores, possibilidade de refazer os microdados a partir das tabelas simplificadas e programas originais	Mitigar	5	5	25	Sim
R03 – Falha/erro de processamento	Procedimento de tratamento de erros previsto no Guia de Procedimentos para Tratamento de Erros de Divulgação do INEP (para reduzir o impacto), possibilidade de refazer os microdados a partir das tabelas simplificadas e programas corrigidos.	Reduzir	10	5	50	Sim
R04 – Reidentificação de dados pseudonimizados.	Adoção do modelo simplificado de microdados	Reduzir	5	10	50	Sim
	Criação de medidas alternativas para acesso aos dados (ex: Serviço de Acesso a Dados Protegidos)	Reduzir	5	10	50	Sim

Legenda: P – Probabilidade; I – Impacto. Aplicam-se as mesmas definições de Probabilidade e Impacto da seção 6.

¹ Efeito resultante do tratamento do risco com a aplicação da(s) medida(s) descrita(s) na tabela. As seguintes opções podem ser selecionadas: Reduzir, Evitar, Compartilhar e Aceitar.

² Risco residual é o risco que ainda permanece mesmo após a aplicação de medidas para tratar o risco.

³ Medida aprovada pelo controlador dos dados pessoais. Preencher a coluna com: Sim ou Não.

Fonte: RIPD Relatório de Impacto à Proteção de Dados Pessoais dos Microdados do Exame Nacional do Ensino Médio (ENEM) (SEI 3848205)

5.60. Como se pode observar, as medidas propostas para evitar a reidentificação do titular, quais sejam (i) adoção do modelo simplificado de microdados e (ii) criação de medidas alternativas para acesso aos dados (ex: Serviço de Acesso a Dados Protegidos) mostraram-se suficientes para mitigar o risco identificado, segundo a avaliação do INEP. A mesma conclusão se aplica para os demais riscos identificados.

IV.II RIPD dos microdados dos censos da educação

5.61. Em atenção aos encaminhamentos propostos pela CGF em sua Nota Técnica 46 (SEI nº 3319546), o INEP também apresentou o Relatório de Impacto de Proteção de dados dos Microdados dos Censos da Educação (Relatório RIPD Microdados dos Censos da Educação, SEI nº 3848206). É esse documento que se passa a analisar no presente tópico, seguindo o roteiro da seção anterior, em conformidade as orientações publicadas pela ANPD em seu sítio eletrônico.

5.62. Em relação ao primeiro tópico (a), observa-se que o RIPD novamente indicou o próprio INEP como controlador e operador de dados, sendo que a figura de operador de dados foi atribuída às Diretorias de Estatísticas Educacionais (DEED) e de Tecnologia e Disseminação de Informações Educacionais (DTDIE). Dessa forma, no presente caso, o INEP não somente toma as decisões referentes ao tratamento de dados pessoais (controlador), mas também é aquele que realiza o tratamento de dados pessoais (operador), por meio de diretorias específicas. Verifica-se, também, que o RIPD indicou corretamente o encarregado, sua identificação e informações de contato (e-mail e telefone).

5.64. No tocante ao item (b) outras partes interessadas/envolvidas, assim como no caso do RIPD do ENEM, o INEP mencionou o envolvimento da própria ANPD, por meio de reunião e também de manifestação nos autos do presente processo, a partir de orientações e encaminhamentos propostos na Nota Técnica 46 (SEI nº 3319546). Ademais, também foram endereçadas no RIPD as preocupações levantadas pela CGU, por meio de Nota Técnica, acerca de aspectos gerais envolvidos na publicação dos microdados dos censos escolares.

5.68. Como justificativa da necessidade (c), o relatório aponta a solicitação da ANPD por meio da Nota Técnica 46, enquadrando-se na hipótese do art. 38 da LGPD. Muito embora tal justificativa seja válida e suficiente, importa registrar, a título de esclarecimento, que cabe ao controlador gerir os riscos associados a todo tratamento de dados pessoais que realiza. Considerando-se que o RIPD abriga "a análise do controlador com relação a medidas, salvaguardas e mecanismos de mitigação de risco adotados", importa chamar atenção para o constante na página da ANPD sobre quando é recomendável pela ANPD a elaboração do RIPD conforme já abordado no [item 5.43](#).

5.70. No item (d), o INEP aponta como processos que justificam a confecção do RIPD os processos institucionais finalísticos das pesquisas estatísticas no Inep envolvidos no tratamento de dados pessoais do Censo Escolar da Educação Básica e do Censo da Educação

Superior.

5.72. Como sistema de informação relacionado ao processo de elaboração e divulgação dos microdados dos censos educacionais (e), o RIPD menciona que o INEP utiliza soluções próprias e específicas desenvolvidas pelo Instituto. Para o Censo Escolar da Educação Básica, o Inep utiliza o Sistema Educacenso, e para o Censo da Educação Superior, o Instituto utiliza o Sistema Censup.

5.74. O relatório também descreve como é realizado o tratamento dos dados pessoais no âmbito dos censos escolares (f). De acordo com o RIPD, o referido tratamento envolve a coleta, produção, recepção, classificação, utilização, acesso, reprodução, processamento, arquivamento, armazenamento, avaliação e controle da informação, difusão e extração de dados.

5.75. Os dados pessoais são coletados a partir de um instrumento estruturado de pesquisa (um questionário), por meio eletrônico, com frequência anual e referência a um período específico, a cada edição das pesquisas, por meio de resposta direta em um sistema eletrônico próprio, disponível via Internet, de acesso controlado, ou por meio de um processo de migração de dados.

5.76. Os dados pessoais tratados pela pesquisa têm por origem os registros administrativos das instituições e redes de ensino. Essas instituições detêm as informações de interesse para as pesquisas do INEP, e são tratadas e transpostas conforme a metodologia estabelecida pelo Instituto, por meio da ação de um usuário do sistema - um informante da pesquisa -, que não necessariamente é o titular dos dados pessoais. O INEP coordena o processo, estabelece o instrumento, os prazos, a metodologia de coleta, os sistemas eletrônicos de informação. Assim, as escolas e instituições de ensino superior, por meio de agentes públicos próprios, reportam os dados requeridos, tendo por referência os registros administrativos e acadêmicos disponíveis. Os dados são recebidos e tratados pelo Instituto, e subsidiam processos de controle de qualidade e sínteses preliminares que são compartilhadas com os agentes públicos dessas instituições, que os avaliam, ratificam e retificam conforme necessidade, concluindo sua participação na coleta.

5.77. Os dados oriundos dos sistemas eletrônicos de coleta e tratamento de dados formam uma base denominada de “produção”, acessível apenas por meio dos sistemas internos, e por servidores públicos ou colaboradores do INEP que atuam especificamente no tratamento de dados para compilação em uma base estatística. Essa segunda base de dados (estatística) é pseudomizada, ou seja, não contém os dados pessoais de identificação. O RIPD destaca que o acesso a essa base de dados somente é possível a partir da infraestrutura interna do órgão, em ambiente específico, e para servidores e colaboradores que atuam no desenvolvimento de produtos de disseminação estatística, estudos e novas pesquisas de caráter institucional.

5.78. Os dados pessoais, mesmo após o encerramento de cada edição das pesquisas, são mantidos pelo Instituto, conforme prevê a legislação, para dar sustentação ao desenvolvimento de novos estudos e pesquisas, pelo próprio órgão e/ou outros pesquisadores da Academia e da Administração Pública. Segundo o INEP, sempre é observado o requisito de preservação da privacidade das pessoas retratadas nas respectivas pesquisas, por processos de anonimização, quando possível; mas também por procedimentos de controle de acesso e extração (controle de divulgação estatística), assim como pelo uso de tecnologias de segurança da informação.

5.79. A seguir, são descritos os dados pessoais tratados nos dois Censos Educacionais realizados pelo INEP.

5.80. O Censo Escolar da Educação Básica trata dados pessoais de quatro titulares diferentes: (i) o gestor escolar; (ii) o profissional escolar em sala de aula; (iii) o aluno e (iv) o usuário do Sistema Educacenso. Abaixo estão indicados os dados tratados de cada um desses titulares.

- Gestor escolar: data de nascimento; sexo; nacionalidade; país de nacionalidade (apenas para estrangeiros); UF e município de nascimento; CPF; nome completo dos pais ou responsáveis; maior nível de escolaridade concluído; tipo de ensino médio cursado; dados gerais do curso superior realizado; pós-graduações concluídas; outros cursos específicos realizados (formação continuada com no mínimo 80 horas); cargo; situação funcional/regime de contratação/tipo de vínculo; critério de acesso ao cargo/função; cor/raça; se possui deficiência, transtorno do espectro autista ou altas habilidades/superdotação; tipo de deficiência, transtorno do espectro autista ou altas habilidades/superdotação.
- Profissional em sala de aula: data de nascimento; sexo; nacionalidade; país de nacionalidade (apenas para estrangeiros); UF e município de nascimento; Código do Docente; CPF; nome completo dos pais ou responsáveis; maior nível de escolaridade concluído; tipo de ensino médio cursado; dados gerais do curso superior realizado; formação/complementação pedagógica; pós-graduações concluídas; outros cursos específicos realizados (formação continuada com no mínimo 80 horas); função que exerce na escola; situação funcional/regime de contratação/tipo de vínculo; turma(s) em que atua; áreas/componentes curriculares com os quais atua; cor/raça; se possui deficiência, transtorno do espectro autista ou altas habilidades/superdotação; qual é o tipo de deficiência, transtorno do espectro autista ou altas habilidades/superdotação.
- Aluno: data de nascimento; sexo; nacionalidade; país de nacionalidade (apenas para estrangeiros); UF e Município de nascimento; Código do Aluno; CPF; número de Identificação Social (NIS), carregado da base do Cadastro Único até 2021; número da matrícula da Certidão de Nascimento (certidão nova); nome completo dos pais ou responsáveis; dados de vínculo escolar (turma, tipo de atendimento especializado, transporte escolar); cor/raça; se possui deficiência, transtorno do espectro autista ou altas habilidades/superdotação; tipo de deficiência, transtorno do espectro autista ou altas habilidades/superdotação; recursos para uso em sala de aula e para participação em avaliações do Inep (Saeb).
- Usuário do educacenso (Cadastro Educacenso + Sistema SSI): data de nascimento; sexo; CPF; nome da mãe; cargo; e-mail; endereço de IP e registros de login no sistema (data e hora, localização).

5.81. O Censo da Educação Superior, por sua vez, trata dados pessoais de quatro titulares diferentes: aluno, docente, usuário do Sistema Censup e gestores das IES e mantenedoras. Abaixo estão descritos os dados coletados de cada um desses titulares:

- Aluno: nome, CPF, documento estrangeiro, código do aluno no Inep, código da pessoa física no Inep, data de nascimento, sexo, nacionalidade, país de origem, UF de nascimento, município de nascimento, nome completo da mãe, vínculo a curso de educação superior (código e nome do curso de educação superior, situação do vínculo, registro acadêmico do aluno na IES, polo de apoio presencial no caso de cursos a distância, turno no caso de cursos presenciais, carga horária total do curso, carga horária já integralizada, semestre de conclusão do curso, se é aluno Parfor, se faz segunda licenciatura, se faz formação pedagógica, IES destino no caso de mobilidade acadêmica nacional, país de destino no caso de mobilidade acadêmica internacional, semestre e ano de ingresso no curso, forma de ingresso/seleção realizada para ingresso no curso, tipo de escola que concluiu o ensino médio, tipo de programa de reserva de vagas que possui, tipo de financiamento estudantil não reembolsável que possui, tipo de financiamento estudantil reembolsável que possui, apoio social que possui, atividade extracurricular que realiza, se possui bolsa/remuneração referente a alguma atividade extracurricular que realize, cor/raça e deficiência, transtorno global do desenvolvimento ou altas habilidades/superdotação que possui.
- Docente: nome, CPF, documento estrangeiro, código do docente no Inep, código da pessoa física no Inep, data de

nascimento, sexo, nacionalidade, país de origem, UF de nascimento, município de nascimento, nome completo da mãe, escolaridade, situação do docente na IES (em exercício ou afastado), regime de trabalho, se é docente substituto, se é docente visitante, se estava em exercício em 31/12, atuação do docente na IES, se possui bolsa pesquisa, vínculo do docente a curso, tipo de vínculo do docente visitante (bolsista ou em folha), cor/raça e deficiência, transtorno global do desenvolvimento ou altas habilidades/superdotação que possui.

- Usuário do CENSUP: nome completo, nome social, número de telefone fixo residencial, número celular pessoal, e-mail pessoal, CPF, código do usuário, código da pessoa física no Inep, data de nascimento, identificação das instituições a que está vinculado (código, nome e sigla).
- Gestores das IES e mantenedoras: nome completo, número de telefone, e-mail, CPF e cargo que ocupa (dirigente principal, procurador institucional ou representante legal).

5.82. Considerando apenas essas duas pesquisas, entre todas as pessoas nelas reportadas, em cada edição anual, o INEP informa tratar uma quantidade substancial de dados de aproximadamente 60 milhões de brasileiros de várias faixas etárias, com predominância do público jovem e em idade ativa. Dessa forma, além de abranger em seu público-alvo majoritariamente crianças e adolescentes, também são tratados dados de docentes e outros profissionais escolares e gestores das instituições de ensino.

5.83. Como dados pessoais sensíveis, observam-se as informações relativas à raça/cor e aquelas concernentes a deficiência, transtorno global do desenvolvimento ou altas habilidades/superdotação.

5.84. Também é importante observar que, em relação aos Censos Educacionais, os titulares dos dados não consentem objetivamente, em instrumento específico, com o tratamento de seus dados para as pesquisas, nem mesmo os informam diretamente ao INEP. A pesquisa se materializa a partir do levantamento de dados indiretos, oriundos dos registros administrativos e escolares, informados por um terceiro, em geral um funcionário da escola ou instituição de ensino. Embora o RIPD não dedique um tópico para abordar a análise da hipótese legal para o tratamento dos dados dos censos escolares (item g), o relatório faz referência à Nota Técnica 46 (SEI nº 3319546) da CGF, em que, nos termos do RIPD, a ANPD "reconheceu que o INEP possui hipótese legal para realizar a divulgação dos microdados." Com efeito, da leitura da referida Nota Técnica da CGF, verifica-se que o tratamento de dados pessoais pelo INEP no âmbito da divulgação dos microdados dos censos escolares está respaldado pela LGPD, pois é desenvolvido no mister do cumprimento de obrigações legais atribuídas à autarquia ou, ainda, para fins da execução da política pública e realização de estudos de acompanhamento do sistema de ensino brasileiro. O tratamento de dados pessoais pelo INEP para a elaboração e posterior divulgação dos microdados dos censos escolares está respaldado pelos inc. II e IV do art. 7º da LGPD. Não há dúvidas, portanto, quanto à legitimidade do tratamento de dados realizado pelo INEP no caso em análise.

5.85. No que tange à finalidade, o INEP ressalta que o tratamento de dados pessoais adotado pelo Instituto por ocasião dos Censos da Educação realiza-se para finalidade estatística, com a produção de pesquisas, avaliações e estudos sobre o sistema educacional brasileiro. Os dados são utilizados em atividade de sustentação de procedimentos técnicos das pesquisas, de maneira a desenvolver a qualidade e fidedignidade dos seus resultados (as estatísticas), bem como para o detalhamento da representação e análise dos resultados para diferentes subgrupos da população pesquisada (coletivos), conforme atributos pessoais dos seus indivíduos (idade, sexo, cor/raça, condição de saúde – deficiência). O INEP acrescenta que o tratamento dos dados pessoais nos Censos Educacionais subsidia, ainda, a integração com outras fontes de informação, ampliando as possibilidades de análises e avaliações, sempre respeitada a finalidade de uso em pesquisa, uma vez que possibilitam que outros estudos sejam desenvolvidos a partir do esforço inicial do tratamento de dados pessoais realizado pelos censos educacionais. O Instituto ainda pondera que suas pesquisas não têm por objetivo a produção e o uso de informação individualizada, quicá pessoal, sobre o titular; apenas a produção de estatísticas sobre os grupos e a população-alvo das pesquisas.

5.87. Acerca da (h) análise de princípios da LGPD, o RIPD foca no aspecto da necessidade e proporcionalidade. O relatório reconhece a importância da transparência e da publicização de dados de interesse público, conquanto reforça a necessidade de preservação do direito dos titulares desses dados à privacidade. Assim, segundo o princípio da necessidade, o tratamento deve ser limitado ao "mínimo necessário para a realização de suas finalidades", abrangendo apenas os "dados pertinentes, proporcionais e não excessivos em relação às finalidades do tratamento de dados" (art. 6º, III LGPD). O princípio da necessidade impõe, portanto, que a coleta (e qualquer outro tratamento) se atenha à menor quantidade possível de dados para o alcance da finalidade proposta.

5.88. Para além da questão do respeito à privacidade, o INEP ressalta que a percepção de insegurança do titular de dados pessoais tratados em pesquisas está relacionada à sua disposição para a colaboração e para a participação esmerada. Assumindo que estatísticas oficiais cumprem uma função pública de grande valor para a sociedade, "torna-se importante apontar que a preservação da identificação do indivíduo representado nas estatísticas não visa somente à garantia do direito da pessoa à intimidade e privacidade, mas sobretudo busca garantir o fundamento sobre o qual a pesquisa se sustenta: a colaboração dos indivíduos em prestar informações fidedignas à pesquisa." De acordo com o Instituto, "a percepção de que os dados pessoais conhecidos pela pesquisa possam ser utilizados para outros fins quebraria o constructo sobre o qual a pesquisa está sustentada: a confiança da sociedade no Instituto de pesquisa, nas garantias legais oferecidas e no próprio levantamento estatístico."

5.89. Note-se, nesse sentido, considerando o impacto que a reidentificação dos titulares pode trazer sobre a própria pesquisa em si, que se sustenta na confiança da confidencialidade, é possível afirmar que qualquer tentativa de reidentificação dos titulares é tratamento posterior incompatível tanto com a finalidade que justificou a coleta quanto com a finalidade que justificou a publicação, consoante se pode depreender da leitura conjunta do artigo 6º, I e com o artigo 7º, §§3º e 7º da LGPD. De fato, excetuada a hipótese prevista no art. 4º, III, da LGPD, qualquer tentativa de reidentificação dos titulares a partir de dados anonimizados divulgados pelo controlador (seja ele público ou privado) cujo tratamento de dados (coleta, processamento e divulgação) tem como premissa a proteção da identidade dos titulares se dá sem amparo nas circunstâncias previstas nos §§3º, 4º e 7º do art. 7º da LGPD porque incompatível com a finalidade que justificou a coleta ou a publicação.

5.90. Tendo em vista que o INEP já possuía evidências de que o evento de reidentificação do titular de dados nos censos não é raro ou fruto de caso fortuito, e sobretudo considerando o resultado obtidos pelo estudo da UFMG, que apontou riscos relevantes à proteção de dados, a providência do INEP foi a retirada dos microdados do site público, na forma até então divulgada, e a substituição do modelo anterior por uma nova estrutura de dados. Segundo alega o Instituto, esse novo modelo mitiga, em grande parte, a abrangência dos riscos identificados. Além disso, o INEP vem promovendo "a ampliação da disseminação dos resultados da pesquisa em outros canais, capazes de tratar diferentes necessidades de uso da informação por diferentes perfis de interessados, sempre preservando o compromisso institucional com o sigilo da informação pessoal coletada para fins de pesquisa."

5.91. Convém registrar que, diferentemente do RIPD dos microdados do ENEM, no RIPD dos censos educacionais o INEP não indicou quais alterações foram implementadas ao nível das variáveis, mas apenas destacou que a elaboração do novo modelo "pode incluir técnicas para a agregação de dados para níveis menos detalhados, generalização de categorias de variáveis, supressão de informações pessoais e que possibilitem a identificação de pessoas, divulgação de bases de dados amostrais, adoção de metodologia de tratamento por privacidade diferencial, esse último ainda carecendo de uma aplicação de consenso."

5.92. No que tange ao tópico i) riscos identificados ao titular, o INEP ressalva que, para efeito do relatório em análise, o escopo da

avaliação de riscos limitou-se ao evento “reidentificação do titular de dados pessoais do público-alvo da pesquisa dos Censos da Educação no arquivo de microdados públicos divulgado do Inep”. O Instituto argumenta que foi esse o evento que motivou a alteração de estrutura de um produto de disseminação estatística - o arquivo de microdados públicos da pesquisa. O INEP ressalta, ainda, que foi esse o objeto da pesquisa realizada pelo Departamento de Ciências da Computação da UFMG.

5.93. No que tange ao estudo que embasou as alterações realizadas, foi possível concluir que a reidentificação indevida do titular é um evento com alta probabilidade de ocorrência, como pode ser observado nos resultados sumarizados abaixo:

No caso do Censo da Educação Superior de 2018:

- uma combinação de 3 quaseidentificadores (dia e ano de nascimento, e código do curso) permite ao adversário reidentificar com absoluta certeza até 39% dos indivíduos na base, ou aproximadamente 4.200.000 estudantes;
- uma combinação de 4 quaseidentificadores (dia, mês e ano de nascimento, e código do curso) pode levar à reidentificação de até 80% dos indivíduos, ou aproximadamente 8.600.000 estudantes;
- na mesma base a mesma combinação de 3 quaseidentificadores permite ao adversário reidentificar um indivíduo aleatoriamente selecionado como alvo com probabilidade de até 56% após a realização do ataque; enquanto a mesma combinação de 4 quaseidentificadores eleva essa probabilidade a até 87%.

No caso do Censo da Educação Básica de 2019:

- uma combinação de 3 quaseidentificadores (mês e ano de nascimento e código da escola em que estudo) permite ao adversário reidentificar um indivíduo aleatoriamente selecionado como alvo com probabilidade de até 29,64% após a realização do ataque;
- enquanto uma combinação de 4 quaseidentificadores (mês e ano de nascimento, município de nascimento e código da escola em que estudo) pode elevar essa chance de sucesso a até 49,86%;
- já o uso de todos os 10 quaseidentificadores (mês e ano de nascimento, município de nascimento e endereço, nacionalidade e país de origem, sexo, cor/raça, código da escola e dependência administrativa da escola) eleva o risco a 75,51%.

5.94. Acrescente-se que o estudo também apontou que a disponibilização dos arquivos em série histórica aumenta a probabilidade de reidentificação, uma vez que oferece mais características sobre a mesma pessoa ao longo do tempo.

5.95. Diante da constatação de elevado risco de reidentificação do titular, o RIPD elenca algumas possibilidades de uso indevido das bases publicadas, incoerentes com a finalidade para a qual os dados foram originalmente coletados e tratados:

- perseguição e atendido contra a vítima: Por exemplo, um *stalker* poderia usar da informação para aumentar a sua chance de sucesso em caso de intentar alguma abordagem física da vítima;
- enriquecimento de outras bases de dados e cadastros, de domínio particular e/ou público para finalidades diversas, agregando-se a elas interesse e valor não compatíveis com a expectativa inicial daquelas bases de dados;
- uso de informação pessoal para estelionato e golpes, complementando informações requeridas em processos administrativos e/ou judiciais ou servindo como validadores de competência do interlocutor junto ao titular em confirmação de situações pessoais;
- aplicações em decisões automatizadas para excluir determinadas pessoas do quadro de clientes, de um novo serviço, de processos seletivos diversos, de quadro de beneficiários, de processo de contratação de colaboradores, de campanhas publicitárias direcionadas; – perfilização de hábitos, comportamentos, consumo, situação econômica, situação de saúde, preferências e interesses pessoais.

5.96. Ademais, o RIPD destaca que, entre os indivíduos cujos dados integram a base, encontram-se pessoas protegidas pela justiça, refugiados e outras em variadas situações de vulnerabilidade. Acrescenta que, ao longo do tempo, a pesquisa tende a compilar dados de toda a população nacional, pois possui cobertura nacional e grande abrangência. Essa condição, associada à alta probabilidade de identificação observada no estudo realizado, culminaria por sujeitar um número expressivo cidadãos a um contexto de fragilidade, tornando relevantes mesmo impactos individuais pequenos, diante da abrangência e do horizonte temporal das pesquisas.

5.97. O quadro abaixo resume os resultados alcançados pelo INEP após a aplicação das medidas citadas para mitigar o risco de reidentificação dos titulares de dados pessoais nos microdados públicos dos Censos da Educação no formato antigo. Registre-se que tal evento foi inicialmente classificado, tanto em relação à probabilidade quanto ao impacto, em nível elevado.

Quadro 3 - Medidas para tratar o risco

Risco	Medida(s)	Efeito sobre o Risco	Risco Residual		Medida(s) Aprovada(s)
			P	I	
Reidentificação do titular de dados pessoais nos arquivos públicos de microdados divulgados no site do Inep.	1. Alteração da estrutura dos arquivos.	Reduzir	Média	Médio	Sim
	2. Acesso controlado por meio do Sedap.		Baixa	Baixo	Sim

Fonte: Relatório de Impacto à Proteção de Dados Pessoais dos Microdados dos Censos da Educação (SEI 3848206)

5.98. Como se pode observar, a nova abordagem adotada, considerando como medidas (i) a alteração da estrutura dos arquivos e (ii) o acesso controlado por meio do SEDAP, segundo a avaliação do INEP, possibilitou uma redução do risco identificado, levando à aprovação das medidas propostas.

5.99. Por fim, o RIPD enfatiza que o INEP segue estudando e desenvolvendo estratégias de otimização das medidas adotadas, e reconhece que "há espaço para o desenvolvimento de novas medidas, novos produtos e canais de atendimento, que podem contribuir para aumentar a utilidade dos produtos e serviços de disseminação estatística sem comprometer a privacidade das pessoas."

6. CONCLUSÃO

6.1. Conforme observado ao longo da presente análise, o INEP está amparado legalmente para realizar o tratamento de dados consistente com a elaboração e a divulgação de microdados relativos ao exame do ENEM e aos censos educacionais. No entanto, ao publicá-los, o Instituto não pode fazê-lo sem a observância dos requisitos e preceitos estabelecidos pela LGPD.

6.2. Como visto, a anonimização nem sempre é a medida de segurança mais adequada e suficiente a ser adotada no contexto de publicação de dados pessoais. Em realidade, a principal orientação da LGPD consiste na necessidade de avaliação de riscos e de adoção de

medidas proporcionais para mitigar a ocorrência de danos. Por esta razão, a possibilidade de eventual identificação dos titulares, ou a admissão de algum grau de risco de identificação, no âmbito do cumprimento de determinações legais ou regulamentares, são compatíveis com a LGPD, desde que adotadas as salvaguardas apropriadas.

6.3. Com a elaboração dos RIPDs solicitados pela ANPD no escopo do presente processo de fiscalização, o INEP foi capaz de descrever os processos de tratamento de dados pessoais que podem gerar riscos às liberdades civis e aos direitos fundamentais, bem como indicou medidas, salvaguardas e mecanismos para mitigação de riscos identificados. Igualmente justificou com base em dispositivos legais e regulamentares, porque o anonimato é uma premissa para a divulgação dos dados do ENEM e do Censo Educacional.

6.4. Segundo a análise, o principal risco mapeado consiste na possibilidade relevante de reidentificação do titular a partir da base de microdados publicada na metodologia antiga. Essa possibilidade concreta significa não somente uma violação ao direito constitucional do cidadão à intimidade e privacidade, mas também representa uma ameaça ao fundamento sobre o qual a pesquisa se sustenta: a colaboração dos indivíduos em prestar informações fidedignas à pesquisa.

6.5. O risco de reidentificação torna-se ainda mais grave tendo em vista que a base de dados no censo escolar do ensino médio contempla crianças e adolescentes em idade escolar obrigatória - dos 4 aos 17 anos. Convém lembrar que, acompanhando a lógica protetiva do ECA (Estatuto da Criança e do Adolescente), a LGPD, em seu art. 14, determina que dados pessoais de crianças e de adolescentes deverão ser tratados em seu melhor interesse, e ainda acrescenta requisitos específicos para o tratamento de dados desse público mais vulnerável, o que denota uma preocupação adicional do legislador.

6.6. Diante dos riscos identificados, o INEP propôs a adoção de duas medidas principais. A primeira consiste na modificação da estrutura dos arquivos que serão publicados, por meio de técnicas para a agregação de dados para níveis menos detalhados, generalização de categorias de variáveis, supressão de informações pessoais e que possibilitem a identificação de pessoas, divulgação de bases de dados amostrais, adoção de metodologia de tratamento por privacidade diferencial. A segunda refere-se ao controle de acesso, mediante procedimento específico requerido, aos pesquisadores externos, mas apenas por meio do Serviço de Acesso a Dados Protegidos do INEP e para finalidade coerente com aquela que ensejou o tratamento inicial dos dados pessoais, qual seja: a produção de estudos e pesquisas, garantida a não identificação do titular de dados pessoais. Tais medidas foram consideradas suficientes pelo INEP para mitigar os riscos mapeados.

6.7. **Por todo o exposto, a CGF conclui que o INEP endereçou adequadamente as preocupações levantadas pela coordenação na Nota Técnica 46, e considera atendida a determinação da ANPD de elaboração do Relatório de Impacto à Proteção de Dados para fins adequação da divulgação dos microdados do censo escolar e do Enem à LGPD.**

6.8. Outrossim, a CGF recomenda ao INEP a revisão contínua dos RIPDs apresentados, em especial, quando houver fatos novos que possam ensejar mudanças nos riscos identificados, tais como alteração nas operações de tratamento, identificação de novos fatores de risco, agravamento dos fatores de risco anteriormente identificados, ou em caso de novas regulamentações ou orientações emitidas pela ANPD. Recomenda-se, ainda, a expansão e o aprimoramento dos canais de acesso controlado às bases de microdados destinados aos pesquisadores externos e demais cadastrados.

6.9. Por fim, é importante ressaltar que a análise desenvolvida na presente Nota Técnica restringiu-se à atual metodologia de divulgação dos microdados dos censos escolares e do Enem face à LGPD e aos RIPDs correspondentes. Isso não significa, necessariamente, que a metodologia adotada anteriormente seria inadequada do ponto de vista da LGPD. Qualquer conclusão nesse sentido demandaria uma análise distinta, considerando as particularidades de tal metodologia, bem como os RIPDs associados, o que não foi feito para o presente caso.

6.10. Acrescente-se, ainda, que a presente análise não significa um engessamento sobre a forma como os microdados são divulgados atualmente. O INEP possui, dentro de suas competências e discricionariedade, liberdade para alterar a forma de divulgação dos microdados, devendo, apenas, verificar previamente quais seriam os impactos para os titulares de dados observando-se as diretrizes da LGPD.

7. ENCAMINHAMENTOS

7.1. Sugere-se o encaminhamento da presente nota técnica ao INEP para ciência, colocando-se a ANPD à disposição para maiores esclarecimentos, se necessários. Adicionalmente, sugere-se ao INEP a publicação dos RIPDs analisados, no que couber, como demonstração de conformidade do Instituto com a segurança dos dados pessoais que estão sob sua responsabilidade e o seu compromisso com a privacidade dos titulares, além de atender aos princípios do livre acesso, da transparência e da responsabilização e prestação de contas, previstos, respectivamente, pelo art. 6º, incisos IV, VI e X, da LGPD.

7.2. Considerando, ainda, a intersecção do tema ora analisado com matéria de competência da Controladoria-Geral da União - CGU, sugere-se o encaminhamento de cópia da versão pública desta Nota Técnica à CGU.

7.3. Tendo em vista as conclusões descritas nesta Nota Técnica, sugere-se também o arquivamento do presente feito.

7.4. À consideração superior.

ULLIANA CERVIGNI MARTINELLI

Especialista em Políticas Públicas e Gestão Governamental

RAVVI AUGUSTO DE ABREU C. MADRUGA

Coordenador de Fiscalização

[1] Esta Nota Técnica foi elaborada com a colaboração de Raíssa Alencar de Sá Barbosa, quando servidora nesta Coordenação-Geral de Fiscalização.

[2] BRASIL. Decreto nº 6.425/2008. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/decreto/d6425.htm. Acessado em 08 ago 2023.

[3] Ministério da Educação. Portaria nº 807, de 18 de junho de 2010. Disponível em: <https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?journal=1&pagina=71&data=21/06/2010>. Acessado em 08 ago 2023.

[4] O INEP foi transformado em Autarquia federal vinculada ao Ministério da Educação (MEC) pela Lei nº 9.448/1997 e suas finalidades estão descritas no art. 1º dessa lei, compreendendo a organização e a manutenção do sistema de informações e estatísticas educacionais (I); o planejamento, a orientação e a coordenação do desenvolvimento de sistemas e projetos de avaliação educacional, visando o estabelecimento de indicadores de desempenho das atividades de ensino no País (II); a prestação de subsídio à formulação de políticas na área de educação, mediante a elaboração de diagnósticos e recomendações decorrentes da avaliação da educação básica e superior (V); a promoção da disseminação de informações sobre avaliação da educação básica e superior (VIII).

[5] WIMMER, Miriam. "O regime jurídico do tratamento de dados pessoais pelo Poder Público". In: DONEDA, Danilo et. Al. **Tratado de**

proteção de dados pessoais. Rio de Janeiro: Forense, 2021. p. 276.

[6] GOMES, Maria Cecília Oliveira. "Relatório de impacto à proteção de dados pessoais". In: Revista do Advogado. São Paulo: Associação dos Advogados de São Paulo, 2019. p. 178.

[7] https://www.gov.br/anpd/pt-br/canais_atendimento/agente-de-tratamento/relatorio-de-impacto-a-protecao-de-dados-pessoais-ripd/relatorio-de-impacto-a-protecao-de-dados-pessoais#p1. Acessado em 22 jun 2023.

[8] Acerca da indicação da Diretoria de Avaliação de Educação Básica, unidade administrativa interna do INEP, convém destacar que a designação de operador não deveria ser entendida como uma norma de distribuição interna de competências e responsabilidades. É nesse sentido que consta no Guia de Agentes de Tratamento e Encarregado (p. 18, https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/guia_agentes_de_tratamento_e_encarregado_defeso_eleitoral.pdf): "... o operador será sempre uma pessoa distinta do controlador; isto é, que não atua como profissional subordinado a este ou como membro de seus órgãos." Ainda que o guia se refira expressamente a pessoas naturais (empregados, administradores, sócios, servidores e outras pessoas naturais que integram a pessoa jurídica), a orientação também é válida para unidades administrativas que fazem parte do controlador. Em síntese, só há sentido em nomear/indicar um operador quando parte das atividades do tratamento é desempenhada por outra pessoa que não o controlador. Nas situações em que o controlador realiza o tratamento integralmente, a designação de um operador é desnecessária.

[9] https://www.gov.br/anpd/pt-br/canais_atendimento/agente-de-tratamento/relatorio-de-impacto-a-protecao-de-dados-pessoais-ripd/relatorio-de-impacto-a-protecao-de-dados-pessoais#p3. Acessado em 09 ago 2023.



Documento assinado eletronicamente por **Ulliana Cervigni Martinelli, Especialista em Políticas Públicas e Gestão Governamental**, em 25/09/2023, às 11:53, 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 **Ravvi Augusto de Abreu Coutinho Madruga, Coordenador(a)**, em 25/09/2023, às 12:09, 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 do documento pode ser conferida informando o código verificador **4038352** e o código CRC **45117D60** no site: https://super.presidencia.gov.br/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0