



INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA

NOTA TÉCNICA Nº 14/2021/CGIM/DAEB

PROCESSO Nº 23036.004635/2021-15

1. ASSUNTO

1.1. Esta nota técnica tem por objetivo fundamentar a necessidade de modificações urgentes nos microdados do Exame Nacional do Ensino Médio (ENEM) divulgados anualmente no portal do INEP, de modo a atender as exigências trazidas pela recente Lei Geral de Proteção de Dados (LGPD), incorporada ao ordenamento jurídico brasileiro por meio da Lei nº 13.709, de 14 de agosto de 2018.

2. REFERÊNCIAS

2.1. BRASIL. Lei nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9394.htm>. Acesso em 19 jul. 2021.

2.2. BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm>. Acesso em 13 jul. 2021.

2.3. TEIXEIRA, Tarcísio; ARMELIN, Ruth Maria Guerreiro da Fonseca. **Lei geral de proteção de dados pessoais**: comentada artigo por artigo. Salvador: Editora Juspodvium, 2019.

3. SUMÁRIO EXECUTIVO

3.1. O INEP precisa se adaptar à nova Lei Geral de Proteção de Dados Pessoais (LGPD), uma vez que várias atividades desempenhadas corriqueiramente pelo instituto estão contidas no amplo conceito de tratamento de dados pessoais previsto na legislação.

3.2. Os microdados da edição de 2020 do ENEM encontram-se prontos para divulgação, seguindo o mesmo modelo utilizado em edições anteriores. A princípio são dados anonimizados, uma vez que não consta na base pública qualquer variável que permita a identificação direta do participante. E, sendo dado anonimizado, estaria fora de escopo da recente legislação de proteção de dados pessoais. Ocorre que a LGPD não considera como anonimizado o dado que possa ser identificado utilizando meios técnicos razoáveis e disponíveis na ocasião de seu tratamento.

3.3. Entretanto, não há clareza sobre o que seriam meios técnicos razoáveis de tratamento, de modo que a CGIM não tem parâmetros técnicos para assegurar que os microdados públicos do ENEM e dos demais exames e avaliações são, conforme conceito estabelecido na lei, dados anonimizados. Com a vigência da LGPD e a ausência de orientações práticas sobre o seu cumprimento, a CGIM optou por não divulgar os microdados do ENEM 2020 no modelo tradicionalmente utilizado para as edições anteriores.

3.4. É fundamental a realização de um estudo para avaliar os produtos informacionais produzidos pela DAEB (microdados e sinopses estatísticas), gerando indicadores de risco objetivos para que seja possível avaliar com segurança, se os mesmos cumprem os requisitos previstos na legislação de proteção de dados pessoais, notadamente a LGPD. Enquanto isso não for possível, a CGIM sugere a adoção de um modelo simplificado de microdados, eliminando da base pública as variáveis que facilitem a identificação indevida do participante. Trata-se de uma sugestão, mas que necessita do respaldo institucional da DAEB, da PROJUR e da Presidência do INEP. Uma vez aprovada, a CGIM encaminhará a publicação dos microdados do ENEM 2020 utilizando o modelo simplificado proposto.

3.5. A CGIM também precisa de orientações sobre como proceder com relação aos microdados do ENEM das edições anteriores (1998 a 2019) e dos demais exames e avaliações, atualmente disponíveis

para download público no portal do INEP. É preciso decidir se o download desses produtos informacionais deve ser desabilitado temporariamente, enquanto se providencia sua adaptação ao modelo simplificado proposto, caso seja aprovado pela gestão do instituto.

4. ANÁLISE

4.1. A LGPD, vigente desde agosto de 2020, promulgou o sistema protetivo dos dados pessoais no direito brasileiro, estabelecendo princípios norteadores da coleta, do compartilhamento e do tratamento de dados pessoais, além de um conjunto de obrigações aos controladores e responsáveis por essas atividades. A lei brasileira foi fortemente inspirada pela legislação europeia, o Regulamento Geral de Proteção de Dados (GDPR), que foi a sua grande propulsora. O projeto de lei tramitava já há algum tempo no Congresso Nacional, mas o tema passou a ser tratado com maior urgência diante do receio de operações comerciais serem barradas devido à falta de equivalência entre as normas protetivas da União Europeia e do Brasil.

4.2. O tema da proteção de dados pessoais é bastante complexo e sensível, pois depende de um amplo conjunto de conceitos, sujeitos à interpretação, bem como por conta dos inúmeros interesses envolvidos. Importante destacar que o objetivo da legislação é tutelar o indivíduo titular dos dados pessoais, e não os dados em si. A proteção estatal se faz necessária diante dos potenciais riscos que o tratamento de dados pode causar ao indivíduo, que poderá ser afetado em sua privacidade, caso os limites legais não sejam resguardados. É fundamental manter o equilíbrio entre a individualidade e a privacidade das pessoas com a livre iniciativa comercial e de comunicação (TEIXEIRA; ARMELIN, 2019).

4.3. Todas as instituições, públicas e privadas, deverão se adaptar à lei. E o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) não é exceção. Dentre as diversas atividades desempenhadas pelo instituto, várias delas estão contidas no amplo conceito de tratamento de dados pessoais previsto no art. 5º, IX 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*".

4.4. Os participantes dos exames e avaliações educacionais operacionalizados pelo INEP informam seus dados pessoais com a expectativa de que a informação seja tratada exclusivamente para a finalidade a que se propõe, assegurando sua privacidade. Nesse contexto, cabe ao instituto demonstrar que os dados por ele tratados não permitem a identificação ou re-identificação desses indivíduos, requerendo, por vezes, a atualização dos procedimentos técnicos e dos produtos de divulgação estatística. Ao mesmo tempo, deve buscar desenvolver novos produtos informacionais para atender demandas cada vez mais recorrentes da sociedade e dos governos por novas informações e prazos mais curtos de atualização.

4.5. Recentemente, por meio da Portaria nº 592, de 10 de novembro de 2020 (Documento SEI 0602249), instituiu-se uma Força-Tarefa para diagnóstico dos impactos da LGPD no âmbito do INEP. Entretanto, suas atividades estão suspensas e nenhum parecer ou orientação prática foi emitida, com o objetivo de guiar as áreas técnicas na adequação aos requisitos legais dos produtos informacionais (tais como microdados e sinopses estatísticas) por elas produzidos.

4.6. O Exame Nacional do Ensino Médio (ENEM) foi criado em 1998 com o objetivo de avaliar o desempenho do estudante ao fim da escolaridade básica. A participação no exame é voluntária e dele podem participar alunos que estão concluindo (concluintes), os que já concluíram em anos anteriores (egressos) ou aqueles que ainda estão cursando o ensino médio (treineiros). Em 2009 ocorreu a reformulação metodológica do ENEM, que passou a ser utilizado para: certificação da Educação de Jovens e Adultos (EJA); critério de seleção para os estudantes que pretendem concorrer a uma bolsa no Programa Universidade para Todos (ProUni); e para estudantes que pretendem concorrer a vagas em algumas instituições federais de ensino superior, acarretando um aumento considerável no número de participantes do exame, bem como a atenção dada a ele pela mídia e pelas instituições de ensino. A partir de 2017, a utilização do ENEM para certificação EJA foi suspensa e substituída pelo fortalecimento do Exame Nacional de Certificação de Competências de Jovens e Adultos (ENCCEJA).

4.7. O INEP divulga anualmente, em seu sítio na internet, os microdados do ENEM. Atualmente, estão disponíveis para download as edições de 1998 até 2019. Os microdados da edição de 2020 encontram-se prontos para divulgação desde de junho de 2021, seguindo o mesmo modelo utilizado nas edições anteriores. Entretanto, com a vigência da LGPD e a ausência de orientações práticas sobre o seu cumprimento, a Coordenação-Geral de Instrumentos e Medidas (CGIM) decidiu por não divulgá-los. Isso porque não há segurança sobre o modelo atual dos microdados do ENEM atender adequadamente os requisitos e vedações previstos na legislação de proteção de dados pessoais. Convém destacar que os dispositivos que preveem sanções em caso de descumprimento já estão em vigor. Essa nota técnica pretende apresentar os motivos dessa insegurança, bem como solicitar orientações por parte da DAEB, da Procuradoria Jurídica (PROJUR) e da Presidência do INEP. A CGIM aguarda um pronunciamento sobre a legalidade da manutenção do modelo atual de microdados, e/ou sobre a necessidade de sua reformulação para o atendimento dos requisitos legais.

4.8. Os microdados se constituem no menor nível de desagregação de dados recolhidos por pesquisas, avaliações e exames realizados. No caso do ENEM, os dados estão por participante. Cabe lembrar que não constam nos dados divulgados quaisquer variáveis que permitam a identificação direta do participante, tais como nome, logradouro do endereço, e-mail, número do CPF, data de nascimento, dentre outras. Inclusive, o próprio número de inscrição real do participante é substituído por uma máscara gerada sequencialmente. Tratam-se, portanto, de dados anonimizados que, a priori, não são considerados dados pessoais e estão fora do escopo da legislação.

4.9. O conceito previsto no art. 5º, III da LGPD estabelece que dado anonimizado é o *“dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento”*. Observa-se, portanto, que a lei não define objetivamente o que são esforços razoáveis, de modo que a equipe técnica não consegue avaliar com segurança se os microdados atendem ao disposto na legislação. Compreende-se que a previsão de um conceito aberto é inevitável, uma vez que as tecnologias de anonimização e reversão evoluem rapidamente, de modo que não haveria como uma lei estabelecer exaustivamente o que seria uma técnica razoável de anonimização de dados (TEIXEIRA; ARMELIN, 2019).

4.10. Complementando, o art. 12, §1º do mesmo diploma legal prevê que *“a determinação do que seja razoável deve levar em consideração fatores objetivos, tais como custo e tempo necessários para reverter o processo de anonimização, de acordo com as tecnologias disponíveis, e a utilização exclusiva de meios próprios”*. Dado o atual estado de desenvolvimento de tecnologia, com a possibilidade de cruzamento de bases, e o filtro por diversas variáveis simultaneamente, nenhum dado pode ser considerado de fato anonimizado. Essa questão preocupa bastante a CGIM.

4.11. O art. 12, §3º da LGPD estabelece o papel da Agência Nacional de Proteção de Dados (ANPD) na definição de padrões e técnicas de anonimização. A agência exercerá a função de reguladora do que se considera razoável em termos de anonimização de dados. Portanto, é imprescindível provocar a ANPD para que se manifeste sobre a necessidade de adequação do formato dos produtos informacionais divulgados pela DAEB, incluindo os microdados do ENEM 2020.

4.12. Em todo caso, imagina-se que a ANPD dependerá de um indicador objetivo a orientar essa decisão. Para isso, é fundamental a realização de um estudo similar ao que foi feito pela Diretoria de Estatísticas Educacionais (DEED) em parceria com a Universidade Federal de Minas Gerais (UFMG). Conforme o Documento de Formalização de Demanda (SEI 0460899), o objetivo da iniciativa foi adequar os microdados dos censos educacionais produzidos pela DEED à legislação vigente, em especial a LGPD, por intermédio dos seguintes objetivos específicos:

- a) Levantar o panorama internacional dos procedimentos de garantia da privacidade em bases de dados de pesquisas estatísticas, em especial nos produtos de disseminação dos microdados;
- b) Alinhar qualquer proposta de adequação aos pressupostos legais e normativos vigentes ou com vigência iminente;
- c) Avaliar o risco potencial de identificação das pessoas a quem os dados estatísticos se referem nos microdados dos censos educacionais;

- d) Apresentar alternativas de métodos de controle de divulgação com base nas abordagens pesquisadas;
- e) Desenvolver e implantar um processo computacional baseado nos métodos de controle de divulgação escolhidos pelo INEP, como forma de aprimorar sua atual política de disseminação de microdados educacionais;
- f) Desenvolver capacidade técnica no Instituto para o tratamento de dados pessoais conforme as disposições da legislação vigente e as melhores práticas internacionais do setor;
- g) Aplicar o conjunto de métodos de controle de divulgação por meio do processo computacional desenvolvido em outro conjunto de microdados dos censos educacionais.

4.13. Foi um amplo estudo que culminou com a entrega de oito produtos, citados individualmente na seção 5 da presente nota técnica. No entendimento da CGIM, deve ser empreendido um estudo similar para avaliar os produtos informacionais produzidos pela DAEB, gerando indicadores de risco objetivos para que seja possível avaliar com segurança, se os mesmos cumprem os requisitos previstos na legislação de proteção de dados pessoais, notadamente a LGPD. Portanto, sem a realização de um estudo dessa natureza, não há como estabelecer com segurança se os microdados do ENEM e dos demais exames e avaliações da educação básica cumprem, de fato, os requisitos previstos na recente legislação de proteção de dados pessoais.

4.14. É importante confirmar se o INEP será caracterizado como órgão de pesquisa, conforme conceito previsto no art. 5º, XVIII da LGPD. Caso o INEP seja assim caracterizado, gozará de prerrogativas que facilitarão o processo de tratamento de dados pessoais. A lei estabelece alguns requisitos para que a instituição seja assim caracterizada. São eles: órgão ou entidade da administração pública direta ou indireta ou pessoa jurídica de direito privado, sem fins lucrativos legalmente constituída sob as leis brasileiras; com sede e foro no país; e que inclua em sua missão institucional ou em seu objetivo social ou estatutário a pesquisa básica ou aplicada de caráter histórico, científico, tecnológico ou estatístico.

4.15. Parece evidente que o INEP atende plenamente os dois primeiros requisitos legais. Entretanto, o termo "*Pesquisa*" não aparece diretamente na missão institucional do INEP. Contudo, aparece o termo "*Produzir conhecimento*", implicitamente associado à realização de pesquisas. É preciso, portanto, definir com clareza se o INEP se enquadra na categoria de órgão de pesquisa prevista na legislação, para que possa usufruir das prerrogativas, apesar de não cumprir, pelo menos explicitamente, um dos requisitos previstos na legislação.

4.16. Outro ponto cuja confirmação é necessária diz respeito à dispensa de consentimento do titular dos dados pessoais, com base nas hipóteses previstas no art. 7º, II, III e IV da LGPD. São elas: cumprimento de obrigação legal ou regulatória, execução de políticas públicas pela Administração Pública e realização de estudos por órgãos de pesquisa (caso o INEP tenha essa condição reconhecida). Não há clareza se o consentimento do titular dos dados pessoais foi dado durante a sua inscrição no exame (mas não abarcaria os participantes das avaliações educacionais), e se esse consentimento abrange a divulgação dos microdados e demais produtos informacionais. Ou se há dispensa na obtenção desse consentimento por se enquadrar em uma das hipóteses legais mencionadas anteriormente.

4.17. As bases de dados mantidas pela DAEB contêm dados pessoais sensíveis dos participantes dos exames e avaliações da educação básica, conforme conceito estabelecido no art. 5º, II da LGPD: "*dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural*". A legislação assegurou para esses dados uma tutela maior (consentimento específico).

4.18. Há nos microdados do ENEM dados pessoais sensíveis tais como: cor/raça, pedidos de atendimento especializado e específico, recursos de atendimento especializado e específico para a realização da prova, etc. Os dados pessoais sensíveis são importantes para a adaptação dos processos de aplicação, assegurando a equidade entre os participantes dos exames e avaliações. Além disso, são

bastante requisitados por pesquisadores e demais usuários dos microdados e, caso deixassem de ser divulgados, trariam impacto negativo às pesquisas desenvolvidas e à imagem institucional do INEP.

4.19. A legislação prevê hipóteses de tratamento de dados pessoais sensíveis, sem a necessidade de consentimento do seu titular, quando for indispensável: para o cumprimento de uma obrigação legal ou regulatória, à execução de políticas públicas pela administração pública ou para a realização de estudos por órgão de pesquisa (caso o INEP tenha essa condição reconhecida), conforme expressamente previsto no art. 11, II, alíneas “a”, “b” e “c” da LGPD. É preciso definir com clareza se o INEP de fato se enquadra nessas hipóteses.

4.20. No que tange às duas primeiras hipóteses previstas nas alíneas “a” e “b” (cumprimento de uma obrigação legal ou regulatória; execução de políticas públicas pela administração pública), é preciso observar ainda o previsto no art. 11, §2º da LGPD, que exige que seja dada publicidade à dispensa de consentimento.

4.21. Também cabe destacar que o art. 11, II, alínea “d” prevê que não será necessário o consentimento do titular para o tratamento de dados pessoais sensíveis, quando for indispensável para o exercício regular de um direito. Talvez essa possibilidade abarque os participantes de exames e avaliações que solicitam atendimento especializado ou específico. Para que possam usufruir desse direito, a coleta de dados sensíveis é indispensável, exonerando o INEP da obrigação de obter o consentimento do participante.

4.22. Contudo, em todos os cenários mencionados anteriormente, é preciso determinar com clareza se a coleta de dados pessoais sensíveis é necessária (art. 6º, III) e adequada (art. 6º, II) para a realização da política pública, de estudos por órgãos de pesquisa e para o exercício regular de um direito pelos seus titulares, ou seja, se atende uma finalidade pré-determinada (art. 6º, I). Trata-se, portanto, dos princípios da finalidade, adequação e necessidade previstos na LGPD (art. 6º, I, II e III): coletar o menor volume de dados pessoais necessário para que se chegue ao fim pretendido de forma adequada. A dispensa do consentimento *“se justifica pelo bem do próprio titular ou da coletividade, desde que a utilização dos dados pessoais sensíveis seja indispensável para isso”* (TEIXEIRA; ARMELIN, 2019, p. 66).

4.23. O art. 14 da LGPD dispõe sobre o tratamento de dados pessoais de crianças e adolescentes. Pelo art. 14, §1º, só poderá ocorrer mediante o consentimento específico e em destaque de pelo menos um dos pais ou responsável legal. Isso significa que se deve buscar por tecnologias capazes de apurar com a máxima precisão se a pessoa que está consentindo com a coleta de dados realmente é um pai ou responsável pelo menor. É insuficiente a simples previsão em termos de uso (TEIXEIRA; ARMELIN, 2019).

4.24. A legislação não prevê hipóteses de dispensa de consentimento nesses casos. Segundo o art. 14, §3º da LGPD, o Poder Público poderá utilizar os dados para contatar pais e/ou responsáveis, desde que os dados sejam utilizados uma única vez e descartados após o uso, o que não é o caso das operações de tratamento de dados realizadas por este instituto. O INEP coleta dados pessoais de crianças e adolescentes para viabilizar a realização de exames e avaliações da educação básica, e divulga essas mesmas informações nos microdados e demais produtos informacionais (de maneira anonimizada, a priori). Caso seja necessário obter o consentimento dos pais ou responsáveis, haverá uma dificuldade operacional que poderia inviabilizar a realização dos exames e avaliações. É preciso esclarecer se o INEP precisará de fato obter esse consentimento específico previsto na legislação.

4.25. Convém destacar que o art. 62 da LGPD prevê que o INEP e ANPD editarão regulamentos específicos para o acesso a dados tratados pela União, para cumprimento do disposto no art. 9º, §2º da Lei nº 9.394, de 20 de dezembro de 1996, a chamada Lei de Diretrizes e Bases da Educação Nacional (LDB). LGPD e LDB determinam a coleta de dados para fins educacionais pela União, incluindo dados pessoais de crianças e adolescentes. Caberá à ANPD regulamentar o acesso para que ambas as leis sejam cumpridas, incluindo a previsão de tratamento desses dados pessoais sem a necessidade de consentimento. É importante esclarecer se a elaboração desse regulamento já foi iniciada, ou quem será o responsável por fazê-lo.

4.26. Evidentemente, a necessidade de consentimento para utilização de dados pessoais em geral, sensíveis e de crianças e adolescentes é relevante para a operação de tratamento de dados como

um todo, Isso inclui a coleta dos dados durante a inscrição do participante no exame ou avaliação, o seu armazenamento na base de dados corporativa, eventuais compartilhamentos dos dados com o consórcio aplicador e outros órgãos públicos, além de inúmeras atividades corriqueiramente realizadas pelo INEP, e que se enquadram no amplo conceito de tratamento de dados pessoais positivado pela LGPD. Essa questão deve ser discutida para resguardar o INEP, com a garantia do cumprimento dos requisitos legais. Entretanto, a divulgação pública de dados anonimizados, como ocorre nos microdados, a priori, estaria fora do escopo da LGPD e não dependeria de consentimento. A questão mais relevante, portanto, pelo menos para o objetivo desta nota técnica, é identificar, com clareza, se os microdados são de fato dados anonimizados, conforme conceito previsto na lei de proteção de dados.

4.27. Ocorre que o conceito de dado anonimizado trazido pela legislação é bastante aberto. Enquanto não houver clareza sobre o alcance desse termo, com a produção de indicadores objetivos para avaliar o risco de re-identificação do participante, o mais indicado é agir de maneira prudente, eliminando dos microdados variáveis que possam suscitar dúvidas sobre sua anonimização.

4.28. Enquanto não se realiza o estudo para avaliar a adequação dos produtos informacionais produzidos pela DAEB, similar ao que foi desenvolvido pela DEED em parceria com a UFMG, a CGIM sugere a adoção de um modelo simplificado de microdados a ser utilizado para a edição 2020 do ENEM, de modo a não atrasar ainda mais a divulgação desses dados, que tradicionalmente ocorre no primeiro semestre de cada ano. De maneira simplificada, o objetivo é eliminar da base pública variáveis que facilitem a identificação indevida do participante. Entretanto, é importante destacar que a CGIM não tem condições técnicas de assegurar que as medidas propostas sejam suficientes para atender os requisitos da LGPD. Por isso, depende da aprovação desse modelo pela DAEB, pela PROJUR e pela Presidência do INEP, assegurando que, ao divulgar os microdados do ENEM 2020 nesse modelo simplificado, a equipe técnica da CGIM não estará descumprindo a LGPD.

4.29. Propõe-se as seguintes adequações nos microdados do ENEM 2020:

4.29.1. **Excluir a variável CO_ESCOLA:**

4.29.1.1. Essa variável corresponde ao código da escola onde o participante afirma ter concluído o ensino médio. O seu preenchimento é autodeclarado pelo participante que se identifica como concluinte do ensino médio, no momento da sua inscrição no ENEM, e sem validação junto ao Censo Escolar. Dessa forma, o fato de o participante ter declarado ser concluinte de determinada escola, não significa que de fato o seja. A validação não ocorre porque, durante o período de inscrições do ENEM, normalmente no primeiro semestre de cada ano, o Censo Escolar ainda se encontra em fase de coleta de dados, inviabilizando esse procedimento. Além disso, essa validação adicional poderia onerar o desempenho do programa de inscrições.

4.29.1.2. Desde que o INEP deixou de calcular e divulgar o ENEM por Escola em 2015, veículos da imprensa utilizam a variável CO_ESCOLA para produzir *rankings* entre as escolas utilizando como base os resultados do ENEM. Essa prática não é recomendada, uma vez que não há garantia da correção do conteúdo da variável utilizada como base para identificar os participantes de cada escola. E além disso, o ENEM não foi concebido para gerar resultados por escola ou rede de ensino.

4.29.1.3. Por conta dessa prática, a CGIM recebe um número considerável de demandas encaminhadas por representantes das escolas, informando que o seu resultado no "ENEM por Escola" está errado, prejudicando a imagem da instituição. Em geral, os demandantes alegam que a média da escola foi calculada de forma incorreta porque foram considerados participantes que não são alunos da escola, e/ou que se deixou de considerar alunos da escola que realizaram o ENEM. Essas demandas são respondidas com a informação de que o INEP não calcula mais o ENEM por Escola desde 2015, que não se responsabiliza por *rankings* elaborados por terceiros e que não recomenda a utilização dos dados do ENEM para essa finalidade.

4.29.1.4. Os argumentos apresentados pelas escolas nessas demandas deixam claro que a informação do código da escola permite, de alguma maneira, a identificação dos participantes por parte dos gestores dessas instituições. Sem essa identificação, a escola não poderia afirmar que foram considerados e/ou desconsiderados indevidamente determinados participantes. Provavelmente isso é possível porque esses indivíduos têm acesso a informações cadastrais detalhadas de seus alunos,

facilitando a localização dos registros correspondentes na base dos microdados por meio de filtros mais complexos envolvendo múltiplas variáveis.

4.29.1.5. Os microdados do ENEM só podem ser disponibilizados publicamente sem consentimento dos titulares dos dados pessoais dos participantes, enquanto contiverem apenas dados anonimizados. Convém lembrar que a LGPD considera anonimizado o dado que não possa ser identificado utilizando meios técnicos razoáveis, conforme previsto no art. 5º, III da lei. Há evidências de que os participantes estão sendo identificados, ainda que não haja clareza quanto à razoabilidade dos meios utilizados.

4.29.1.6. Dessa forma, em fiel observância ao disposto na legislação de proteção de dados pessoais, é prudente excluir dos microdados do ENEM a variável CO_ESCOLA, uma vez que a mesma permite a identificação indevida dos participantes.

4.29.2. Excluir dos microdados informações referentes aos pedidos de atendimento especializado e específico, recursos de atendimento especializado e específico para a realização da prova e cor/raça do participante:

4.29.2.1. As informações referentes à cor/raça do participante e aos pedidos de atendimento especializado e específico e recursos de atendimento especializado e específico para a realização da prova enquadram-se no conceito de dados pessoais sensíveis previsto na LGPD. A legislação estabelece que o tratamento de dados dessa natureza depende de consentimento específico, resguardando algumas hipóteses de dispensa previamente analisadas nesta nota técnica.

4.29.2.2. Em primeiro lugar, convém lembrar que não há clareza com relação ao enquadramento do tipo de tratamento de dados realizado pelo INEP em uma das hipóteses de dispensa de consentimento trazida pelo diploma legal. Evidentemente, essa discussão com relação à necessidade de consentimento não é relevante para os casos envolvendo dados anonimizados.

4.29.2.3. A priori, acredita-se que os microdados do ENEM correspondem a dados anonimizados. Entretanto, como o número de participantes que solicitam atendimento especializado ou específico é muito pequeno se comparado com o número total de inscritos, essas variáveis facilitariam a identificação do participante, de modo que o dado deixaria de ser anonimizado, dependendo de consentimento do seu titular para ser divulgado. Dessa forma, é mais prudente remover essas variáveis dos microdados.

4.29.2.4. Entretanto, espera-se que essa decisão gerará críticas por parte dos pesquisadores e demais usuários dos microdados. Ocorre que o INEP não pode deixar de cumprir determinação legal, ainda que prejudique sua imagem institucional.

4.29.2.5. A variável cor/raça, apesar de ser considerada dado pessoal sensível, pode ser mantida nos microdados no entendimento da CGIM. Acredita-se que a mesma não facilitaria a identificação indevida do participante, por conta do elevado número de participantes em cada uma das suas categorias.

4.29.3. Substituir Idade por Faixa Etária:

4.29.3.1. A variável NU_IDADE disponibiliza a informação da idade do participante no dia 31 de dezembro do ano da edição do ENEM. Essa informação, combinada com outros dados cadastrais eventualmente a disposição dos usuários dos microdados, poderia facilitar a identificação indevida do participante.

4.29.3.2. Para dificultar esse processo, tendo como base a argumentação apresentada para justificar a exclusão da variável CO_ESCOLA, sugere-se que a variável NU_IDADE seja substituída na base dos microdados por uma Faixa Etária, cujas categorias poderiam ser as mesmas empregadas na sinopse estatística do ENEM.

4.29.3.3. Neste caso, não há evidência que comprove a utilização da variável NU_IDADE para identificar os participantes. Entretanto, como a experiência demonstra que a substituição da variável por uma faixa etária não prejudicaria as pesquisas, sugere-se sua realização.

4.29.4. Excluir informações referentes aos municípios de nascimento e residência do participante:

4.29.4.1. 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.

4.29.4.2. 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.

4.29.4.3. Essas informações, combinadas com outros dados cadastrais eventualmente a disposição dos usuários dos microdados, podem facilitar a identificação indevida do participante.

4.29.4.4. Como, a princípio, a experiência demonstra que essas variáveis possuem pouca relevância para os pesquisadores e demais usuários dos microdados, uma vez que a maior parte das pesquisas utilizam como base o município do local de prova (que será mantido nos microdados), sugere-se a exclusão dessas variáveis, dado o risco de identificação dos participantes.

4.29.4.5. A argumentação é a mesma utilizada para justificação a exclusão da variável CO_ESCOLA e a substituição da variável NU_IDADE por uma faixa etária.

4.29.5. Por fim, convém destacar que a CGIM também precisa de orientações sobre como proceder com relação aos microdados do ENEM das edições anteriores (1998 a 2019) e dos demais exames e avaliações, atualmente disponíveis para download público no portal do INEP. Da mesma maneira que para os microdados do ENEM 2020, enquanto não for realizado o estudo para avaliar a adequação dos produtos informacionais produzidos pela DAEB, similar ao que foi desenvolvido pela DEED em parceria com a UFMG, não haverá clareza sobre esses produtos informacionais cumprirem fielmente os requisitos e vedações da LGPD.

4.30. Diante desse cenário, pode-se sugerir que o link para download de todos os produtos informacionais gerados pela DAEB seja desabilitado, até que as adequações reputadas necessárias pelo estudo a ser realizado sejam implementadas e testadas. Evidentemente essa medida seria criticada por pesquisadores e demais usuários dos microdados.

4.31. Como alternativa menos drástica, pode-se realizar adequações similares às propostas para os microdados do ENEM 2020, retirando dados que facilitem a identificação do participante. O problema nesse caso seria a disponibilidade de equipe para viabilizar muitas alterações simultaneamente. A CGIM precisaria planejar-se para executar pontualmente as adequações, e essas atividades competiriam com outras já planejadas. E durante esse período de tempo, os microdados permaneceriam disponíveis para download, potencialmente em descumprimento ao que estabelece a legislação.

5. DOCUMENTOS RELACIONADOS

- 5.1. Portaria nº 592, de 10 de novembro de 2020 (Documento SEI 0602249).
- 5.2. Documento de Formalização de Demanda (SEI 0460899).
- 5.3. Produto nº 01 - Panorama Internacional e o Contexto do INEP (Documento SEI 0546760).
- 5.4. Produto nº 02 - Relatório Revisado (Documento SEI 0571714).
- 5.5. Produto nº 03 - Métodos de Tratamento Aplicáveis (Documento SEI 0596198).
- 5.6. Produto nº 04 - Solução Tecnológica - Documentação (Documento SEI 0653551).
- 5.7. Produto nº 05 - Relatório Técnico Final do Projeto Piloto (Documento SEI 0653553).
- 5.8. Produto nº 06 - Relatório Técnico de Implantação (Documento SEI 0653559).
- 5.9. Produto nº 07 - Atividades de Transferência de Tecnologia (Documento SEI 0676339).
- 5.10. Produto nº 08 - Relatório Final do Projeto (Documento SEI 0678060).

6. CONCLUSÃO

6.1. É imprescindível a realização de um estudo, similar ao realizado pela DEED em parceria com a UFMG, para avaliar a adequação quanto aos requisitos e vedações trazidos pela LGPD dos produtos informacionais (microdados e sinopses estatísticas) elaborados pela DAEB.

6.2. Enquanto esse estudo não for realizado, sugere-se a utilização de um modelo simplificado para os microdados do ENEM 2020, eliminando e/ou substituindo variáveis que facilitem a identificação

indevida dos participantes, assegurando a anonimização dos dados e a dispensa de consentimento do titular para a sua divulgação.

6.3. A CGIM aguarda orientações da DAEB, da PROJUR e da Presidência do INEP sobre a adequação e suficiência das alterações propostas para assegurar o cumprimento da LGPD. Caso as sugestões sejam aprovadas, encaminhar-se-á a publicação dos microdados do ENEM 2020 conforme o modelo simplificado sugerido.

6.4. Também será necessário o recebimento de orientações sobre como proceder com os microdados das edições anteriores do ENEM e dos demais exames e avaliações da educação básica.

6.5. Sugere-se ainda provocar a ANPD para que se manifeste sobre a necessidade de adequação do formato dos produtos informacionais divulgados pela DAEB.



Documento assinado eletronicamente por **Giordano Alan Barbosa Sereno, Servidor Público Federal**, em 28/07/2021, às 11:04, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Joelson Severo dos Santos Azevedo, Coordenador(a) - Geral**, em 28/07/2021, às 11:48, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.inep.gov.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0733088** e o código CRC **D02190E1**.