



NOTA TÉCNICA

MP 954

**RECOMENDAÇÕES PARA O
COMPARTILHAMENTO DE
DADOS ENTRE EMPRESAS DE
TELECOMUNICAÇÃO
E O IBGE**



LAPIN



Sobre esta nota técnica

A edição da **Medida Provisória n. 954**, que dispõe sobre o compartilhamento de dados por empresas de telecomunicações com a Fundação Instituto Brasileiro de Geografia e Estatística (**IBGE**), afeta diretamente o cenário de proteção de dados no Brasil.

A falta de critérios para segurança da informação, a ausência de uma finalidade específica para o tratamento e a coleta desproporcional de dados são alguns dos temas a serem revistos para que se possa permitir o uso das informações de brasileiros para fundamentar o trabalho da fundação.

Com isso em vista, o Laboratório de Políticas Públicas e Internet - **LAPIN**¹ editou esta nota técnica para apresentar suas considerações sobre o tema.

Quem somos nós

O **Laboratório de Políticas Públicas e Internet (LAPIN)** é um *think tank* com sede na capital federal brasileira, de composição multidisciplinar e cujo objetivo apoiar o desenvolvimento de políticas públicas voltadas para a regulação das tecnologias digitais por meio da pesquisa e da conscientização da sociedade.

¹ Este trabalho foi elaborado por: José Renato Laranjeira de Pereira, Paulo Henrique Atta Sarmento, Thiago Guimarães Moraes e Henrique Bawden.

SUMÁRIO

| | |
|--|-----------|
| I - Introdução | 4 |
| II - Finalidade do tratamento | 5 |
| III - Da confusão envolvendo o conceito de dados pessoais | 8 |
| IV - Minimização de dados | 10 |
| V - O período para eliminação dos dados pelo IBGE deve ser vinculado ao estudo estatístico, não ao fim da crise | 13 |
| VI - Medidas adicionais de proteção de dados | 14 |
| VI - Conclusão | 19 |

I - Introdução

Com a proliferação do COVID-19 e as consequentes medidas de isolamento social, o processo de tratamento de dados pela Fundação Instituto Brasileiro de Geografia e Estatística - IBGE para operacionalizar censos foi prejudicado. Diante disto, o Executivo editou a Medida Provisória 954/2020, que dispõe sobre o compartilhamento de dados entre empresas de telecomunicações e o IBGE para possibilitar a realização de “**estatísticas oficiais**” feitas pela fundação.

A norma estabelece que empresas de telecomunicações que trabalham com o Serviço Telefônico Fixo Comutado - STFC e do Serviço Móvel Pessoal - SMP deverão compartilhar dados relativos ao **nome, número de telefone e endereços** de seus clientes, pessoas físicas e jurídicas.

O texto prevê algumas salvaguardas à proteção dos indivíduos cujos dados serão colhidos: o art. 3º estabelece que tais dados teriam caráter sigiloso, com uso exclusivo do IBGE para fim de possibilitar a entrevista não-presencial. Será vedado o seu uso em processos administrativos, fiscais ou judiciais, e a fundação deverá realizar relatório de impacto de proteção de dados pessoais.

Por fim, o art. 4º determina que os dados tratados seriam **deletados ao fim da pandemia**. No entanto, caso houvesse necessidade do IBGE para finalizar estatística oficial, a eliminação poderia ser postergada por até 30 dias.

A MP veda o compartilhamento dos dados pelo IBGE com outros órgãos da administração pública. Por fim, seu art. 2º, §2º, determina que o Presidente do IBGE disporá sobre procedimento para a disponibilização dos dados pelas empresas de telecomunicação.

Por meio da **Instrução Normativa n. 2/2020 do IBGE**, datada do mesmo dia da edição da Medida Provisória, ao invés de detalhar os procedimentos para compartilhamento de dados, apenas sinalizou que “deverão ser disponibilizados no formato e nos veículos de compartilhamento escolhidos pelas empresas de telecomunicações”. Além disso, definiu a Diretoria de Informática da instituição como responsável pela operacionalização da transmissão dos dados. Nenhuma medida adicional de proteção de dados é instituída.

Apesar da necessidade de se produzir dados estatísticos sobre o país, ainda mais em uma situação de crise, há uma série de pontos nos quais **a Medida Provisória é falha** no que tange à proteção dos dados pessoais e da intimidade e privacidade dos indivíduos.

Como será demonstrado a seguir, os textos da MP 954 e da IN/IBGE 2/2020 prevêem a **coleta de dados desnecessários e desproporcionais**, e não estipulam, na ausência da Autoridade Nacional de Proteção de Dados (ANPD), uma auditoria independente para fiscalizar a operação. Além disso, peca ao não estabelecer outras salvaguardas de proteção de dados.

Pelo exposto, o LAPIN se posiciona contrariamente à manutenção da medida provisória nos moldes em que foi editada pelo Presidente da República. Modificações devem ser feitas no que tange à finalidade dos dados, ao modo de coleta e a quais dados serão coletados. Também deve ser endereçada a necessidade de criação de salvaguardas para proteção dos dados dos indivíduos afetados pela medida provisória, conforme será demonstrado.

II - Finalidade do tratamento deve ser melhor especificada

O texto da Medida Provisória estabelece a “*produção de estatística oficial*” como a finalidade do tratamento de dados que será realizado pelo IBGE. No entanto, **não há especificação de qual estatística específica será produzida**, dentre as tantas realizadas pela fundação.

Essa imprecisão na descrição do propósito do tratamento vai contra o art. 6º, Inciso I, da Lei nº 13.709, a Lei Geral de Proteção de Dados (LGPD), que estabelece que a finalidade deve ser específica.² Para que cumpra esse requisito, **é necessário que a MP 954 determine quais estudos estatísticos serão realizados e quais os resultados esperados.**

² Art. 6º As atividades de tratamento de dados pessoais deverão observar a boa-fé e os seguintes princípios:

I - finalidade: realização do tratamento para propósitos legítimos, específicos, explícitos e informados ao titular, sem possibilidade de tratamento posterior de forma incompatível com essas finalidades;

Apesar de não estar expressa na MP, informações extraoficiais indicam que estes dados serão utilizados para a Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC)³. Essa pesquisa envolve a coleta de dados sobre a **raça** dos entrevistados,⁴ o que configura **dados sensíveis**, de acordo com o art. 11, da LGPD.

Esse tipo de tratamento de dados exige mais condições de segurança do que um tratamento de dados pessoais sem categoria específica. Levando em conta que obrigações descritas em leis de proteção de dados são, em grande medida, fórmulas para lidar com riscos aos direitos e liberdades de indivíduos,⁵ procedimentos eficazes de mitigação deve ser colocados em prática.

Isso se aplica principalmente no que diz respeito à criação de maiores restrições de acesso e de garantia de confidencialização dos dados via pseudoanonimização e anonimização, tema que será melhor explorado na **Seção VI**. Em adição a isso, a vedação ao compartilhamento dos dados com outras entidades e a obrigatoriedade de elaboração de relatório de impacto de proteção de dados, já presentes na MP, devem seguidas à risca.

É extremamente importante esses dados sejam coletados pelo IBGE de acordo com a base legal do **consentimento**. Apesar de não ser a única possibilidade jurídica que permite o tratamento de dados sensíveis, é a que permite maior transparência e controle do titular de dados sobre suas informações. Isso não fugiria muito da prática corrente da elaboração da PNADC, que prevê que as pesquisas são feitas de forma **voluntária**.⁶

Essa tem sido a razão pela qual a Comissão Europeia tem recomendado que a coleta de dados sensíveis no bloco regional seja feita a partir de uma cessão

³ SENADO FEDERAL (Brasil). **Operadoras deverão repassar dados de clientes a IBGE para pesquisa por telefone**. [S. l.], 20 abr. 2020. Disponível em: <https://www12.senado.leg.br/noticias/materias/2020/04/20/operadoras-deverao-repassar-dados-de-clientes-a-ibge-para-pesquisa-por-telefone>. Acesso em: 21 abr. 2020.

⁴ INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Manual Básico da Entrevista. Pesquisa Nacional por Amostra de Domicílios Contínua**. Coordenação de Trabalho e Rendimento. Rio de Janeiro: IBGE, 2016. p. 53.

⁵ FINCK, Michèle; PALLAS, Frank. **They who must not be identified—distinguishing personal from non-personal data under the GDPR**. *International Data Privacy Law*, Oxford, p. 25, 10 mar. 2020. Disponível em: <https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf>. Acesso em: 9 abr. 2020.

⁶ IBGE, Op. Cit., p. 38.

voluntária do indivíduo.⁷ Caso isso não fosse seguido, sinalizou que a privacidade de indivíduos seria largamente afetada pela dificuldade em determinar a adequação, necessidade e proporcionalidade da coleta de dados sensíveis obrigatória de indivíduos.

No caso brasileiro, isso se agrava pela ausência de uma entidade como a ANPD em pleno funcionamento para supervisionar os tratamentos realizados. Vale ressaltar que, em contextos de tratamento de dados sensíveis, como é o caso do PNADC, o **consentimento deve ser feito de forma específica e destacada para a finalidade específica**, nos termos do art. 11, I, da LGPD.

Outra questão diz respeito ao recente **convênio firmado entre IBGE e Ministério da Saúde** para usar a PNADC para monitorar a incidência do COVID-19 no Brasil.⁸ Da forma como está redigida, a MP 954 pode permitir a coleta de dados para fundamentar essa pesquisa, o que deve ser proibido caso não haja regulação específica para tratar a matéria.

Com isso, os dados coletados dos brasileiros, tais como seus nomes, telefones e endereços, serviriam para criar correlações com dados de saúde, o que também caracterizaria um tratamento de dados sensíveis. Isso porque dados extraem sua característica de serem sensíveis ou não a partir também do contexto em que são tratados.⁹

A esse respeito, vale trazer uma reflexão a respeito do **conceito de dado pessoal**, especialmente por conta da confusão que tem surgido a respeito do tema. Manifestações de autoridades que revelam a falta generalizada de compreensão sobre esse termo basilar à proteção de dados devem ser endereçadas para melhor esclarecer os pontos trazidos neste trabalho.

⁷ European Commission. **Guidance on Apps supporting the fight against COVID 19 pandemic in relation to data protection**. 17 abr. 2020. Disponível em [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020XC0417\(08\)&from=EN](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020XC0417(08)&from=EN). Acesso em 23 abr. 2020.

⁸ Agência de Notícias IBGE. **IBGE faz parceria com Ministério da Saúde para monitorar casos de Covid-19**, 2 abr 2020. Disponível em <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/27302-ibge-faz-parceria-com-ministerio-da-saude-para-monitorar-casos-de-covid-19>. Acesso em 23 abr 2020.

⁹ ARTICLE 29 WORKING PARTY. Health data in apps and devices. Bruxelas, 2 maio 2015. Disponível em: https://ec.europa.eu/justice/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf. Acesso em: 24 abr. 2020.

III - Da confusão envolvendo o conceito de dados pessoais

Em carta aberta às instituições que entraram com ações diretas de inconstitucionalidade contra a MP 954, **ex-presidentes do IBGE alegaram que** os dados a serem compartilhados “**não incluem informações pessoais**”, e que as “informações são as mesmas que eram antes publicadas nas “páginas amarelas” ou catálogos de telefones”.¹⁰ Essas manifestações são equivocadas, como se verá a seguir.

A disciplina de proteção de dados surgiu na década de 1970, quando o desenvolvimento tecnológico aumentava profundamente a capacidade de processar dados e realizar cruzamentos que permitissem identificar pessoas e extrair inferências sobre elas.¹¹

Nesse contexto, a proteção de dados se instituiu como uma parte da esfera de proteção ao direito à personalidade e liberdade do indivíduo. Por isso, legislações ao redor do mundo têm adotado conceituações amplas de o que seria dado pessoal e, conseqüentemente, de como se daria sua proteção. A tabela abaixo descreve o **conceito** e traz exemplos:

¹⁰ Simon's Site. **Precisamos das estatísticas do IBGE para ajudar a vencer o COVID-19. 20 abr. 2020.** Disponível em <http://www.schwartzman.org.br/sitesimon/?p=6488>. Acesso em 23 abr. 2020.

¹¹ Viktor Mayer-Schönberger. **Generational Development of Data Protection in Europe.** In AGRE, Philip E; Rotenberg, Marc. Technology and Privacy: The New Landscape. Cambridge: The Mit Press. 2001.

| Por que endereço e telefone são dados pessoais? | | |
|---|--------------------------------------|--|
| Conforme a LGPD, art. 5º, I, dado pessoal é a informação relacionada a pessoa natural identificada ou identificável. Este conceito é composto por quatro elementos: ¹² | | |
| Elementos do dado pessoal | Informação | Pode ter natureza objetiva (ex. idade) ou subjetiva (ex. o devedor X é confiável). |
| | Contexto | Um mesmo tipo de dado pode ser considerado pessoal ou não a depender do contexto em que ele é utilizado |
| | Pessoa Natural | Para ser pessoal, a informação deve estar relacionada a um indivíduo humano. |
| | Identificada ou identificável | “Identificada” significa que a ligação ao indivíduo é feita de forma direta, como pelo tratamento de seu nome completo ou sua foto. Como “identificável”, a ligação é indireta, e um processo de cruzamento de dados pode ser necessário para a identificação. Isto contudo não elimina a caracterização do dado como dado pessoal. É o caso de identificadores como o RG, CPF, o endereço e o telefone de uma pessoa natural. |

Cabe conceituar também o que é um **dado pessoal sensível**. Previsto no art. 5º, inciso II, da LGPD, consiste em um dado sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de

¹² Esses elementos são apresentados a partir do conceito legal da LGPD e as referências do Grupo de Trabalho do Artigo 29, Comitê Europeu que regulou matérias de proteção de dados. Para mais informações, checar: Article 29 Working Party, Opinion 4/2007 on the concept of personal data Brussels, 2007. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf. Acesso em: 23 abr. 2020.

caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural.

É vedado realizar o tratamento de dados sensíveis, salvo com base nas exceções previstas no art. 11, da LGPD. Esse tratamento diferenciado ocorre por se tratarem de dados que são capazes de infringir a privacidade ou o exercício de liberdades de seu titular.

A partir disso, qual a diferença entre a coleta de dados pessoais como nome, número e telefone da forma massificada e digital, como propõe a MP 954, e as páginas amarelas? A resposta reside na facilidade de extração de inferências a partir desses dados e a facilidade de pareá-los com os resultados das entrevistas conduzidas pelo IBGE, no atual estágio de desenvolvimento tecnológico.

Dados sobre ideologia, hábitos de consumo, situação de saúde são apenas alguns dos extratos possíveis a partir das informações compartilhadas no escopo da MP. Com isso em mãos, agentes mal intencionados têm a capacidade de influenciar hábitos de consumo, discriminar pessoas com base em sua raça, ideologia ou orientação sexual, recusar crédito e inclusive manipular eleições, como ocorreu com o caso Cambridge Analytica.¹³

IV - Minimização de dados deve permear todo o tratamento

a. Necessidade e adequação

Em seu art. 2º, a MP 954 informa quais dados serão solicitados às empresas de telecomunicação. Isso inclui o nome, o número de telefone e o endereço de seus clientes. Entretanto, esse mesmo artigo, em seu §1º, explicita que esta pesquisa será realizada em “caráter não presencial”.

Para atender ao art. 6º da LGPD, os dados a serem coletados devem ser **necessários e adequados** à finalidade de seu tratamento. Desses dois princípios

¹³ The Guardian. **What is the Cambridge Analytica scandal? - video explainer.** 19 mar 2018. <https://www.theguardian.com/news/video/2018/mar/19/everything-you-need-to-know-about-the-cambridge-analytica-expose-video-explainer>. Acesso em 24 abr 2020.

se subsume a ideia de **minimização**, que determina que a coleta de dados pessoais deve se restringir ao mínimo possível para alcançar o propósito pretendido pelo tratamento.¹⁴

Com isso, surgem as seguintes questões: há necessidade de solicitação de endereço se as pesquisas serão realizadas em “caráter não presencial”? Caso seja necessário para extrair informações estatísticas, não bastaria uma indicação da região ou bairro em que a pessoa reside, e não de seu endereço exato?

Além disso, o nome e o endereço não poderiam ser coletados pelo próprio IBGE durante o contato telefônico e, portanto, evitar o risco de vazamento desses dados durante o processo de compartilhamento de dados entre empresas e fundação?

Para responder a essas questões, cabe compreender o procedimento de realização da PNADC e quais dados são normalmente utilizados pelo IBGE para conduzi-la. Vale mais uma vez trazer que, do texto da MP, **não é possível concluir que os dados serão utilizados só por essa pesquisa**, o que fere o princípio da finalidade disposto na LGPD. Essa inferência se dá com base apenas em informações extraoficiais.

b. Da minimização de dados para realização da Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC)

A PNADC tem o objetivo de produzir informações básicas para o estudo do desenvolvimento socioeconômico do país, bem como permitir a investigação contínua dos indicadores conjunturais de trabalho e rendimento.¹⁵

De acordo com as normas regulamentadoras da pesquisa, os domicílios selecionados para integrar seu Plano Amostral e participar da pesquisa são retirados do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE),¹⁶

¹⁴ Information Commissioner’s Office. **Principle (c): Data minimisation.** Disponível em <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/data-minimisation/>. Acesso em 23 abr. 2020.

¹⁵ IBGE, op. cit., p. 190.

¹⁶ FREITAS, Marcos Paulo Soares de. **Sistema integrado de pesquisas domiciliares: amostra mestra 2010 e amostra da PNAD contínua.** Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2014. p. 14.

uma base de dados pública composta de 78 milhões de endereços urbanos e rurais em todo Brasil.

O plano amostral é feito a partir de uma seleção aleatória de domicílios a serem entrevistados¹⁷ dentro de um certo espaço territorial pré-definido pelo IBGE, de modo a garantir a representatividade de cada região. As pesquisas do PNADC são divulgadas com periodicidade mensal, trimestral e anual.¹⁸

Conforme o IBGE, a cada trimestre, são entrevistados ao todo 211.344 domicílios.¹⁹ Esse número é muito abaixo da quantidade de indivíduos que serão afetados pela **Medida Provisória, que pode alcançar o total de 230 milhões de linhas telefônicas** em uso no Brasil, considerando somente as de celular.²⁰

Isso mostra como é desproporcional a ideia de o IBGE ter acesso aos dados de todos os clientes das empresas de telefonia, sendo que normalmente consegue conduzir suas pesquisas com qualidade utilizando somente os de cerca de 211 mil, em um claro desrespeito ao princípio da minimização de dados.

Por isso é que a coleta desses dados deve ser feita de **modo amostral e aleatório**. Uma solução viável para isso é que **as empresas de telefonia já passem os dados de clientes na quantidade razoável e selecionada de forma aleatória para realização da pesquisa**, nos parâmetros usados nas pesquisas anteriores.

Ultrapassada a questão da quantidade de clientes cujos dados serão compartilhados, cabe agora analisar quais dados de cada indivíduo são necessários para a realização da pesquisa de modo a cumprir com os princípios da necessidade e adequação.

¹⁷ FREITAS, Marcos Paulo Soares de. **Sistema integrado de pesquisas domiciliares:** amostra mestra 2010 e amostra da PNAD contínua. Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2014. p. 22.

¹⁸ ESCOLA NACIONAL DE ADMINISTRAÇÃO PÚBLICA (org.). Censo Demográfico, Pesquisa Nacional por Amostra de Domicílios (PNAD) e PNAD Contínua. Brasília, 26 nov. 2019. Disponível em:

<https://repositorio.enap.gov.br/bitstream/1/4643/6/2019-11-26%20-%20PNAD%2C%20Censo%20Demográfico%2C%20PNADC%20e%20PNS.pdf>. Acesso em: 24 abr. 2020.

¹⁹ IBGE, op. cit., p. 17.

²⁰ Época Negócios. **Brasil tem 230 milhões de smartphones em uso**. 26 abr 2019. Disponível em

<https://epocanegocios.globo.com/Tecnologia/noticia/2019/04/brasil-tem-230-milhoes-de-smartphones-em-uso.html>. Acesso em 22 abr 2020.

Em condições normais, como visto, os agentes do IBGE conduzem a PNADC de forma presencial. Dotados apenas do endereço das pessoas, os contatos são feitos diretamente nas residências, momento em que as demais informações úteis para a pesquisa são coletadas.

Isto posto, **um procedimento a ser seguido pelo IBGE deveria consistir em coletar somente amostragem aleatória do número de telefone dos clientes** com as empresas de telefonia, separados conforme a sua unidade territorial definida pelas normas do PNADC.

Uma vez realizado o contato com cada indivíduo, aí sim a fundação terá a oportunidade de recolher outros dados, tais como nome e endereço do indivíduo. Vale ressaltar que o PNADC é feito de forma voluntária.

Esse procedimento é essencial para garantir maior segurança da informação. Menos informações estarão expostas a vazamentos ou ataques durante a transferência dos dados entre as empresas de telefonia e o IBGE. Além disso, passarão a ser respeitados os princípios da **necessidade** e **adequação** de dados, previstos na LGPD.

V - O período para eliminação dos dados pelo IBGE deve ser vinculado ao estudo estatístico, não ao fim da crise

A medida provisória estabelece, em seu art. 4º, que os dados compartilhados pelas empresas serão deletados após a superação da pandemia. Prevê ainda que este ato ser postergado caso haja comprovada necessidade de uso dos dados para a conclusão de produção estatística oficial pelo IBGE.

Essa redação abre margem para que haja abusos na utilização dos dados, o que viola o princípio da segurança e da finalidade. Afinal, o descarte dos dados não está vinculado ao encerramento do estudo estatístico, mas sim ao fim da pandemia. Tal previsão permite que os dados coletados continuem armazenados por período indeterminado que pode superar em muito o tempo de duração da pesquisa.

Uma vez que a finalidade precípua do IBGE seja a produção de estudos estatísticos, que, por sua natureza, apresentam dados agregados, tão logo esses resultados sejam produzidos e o estudo concluído, os dados pessoais coletados devem ser eliminados.

Por isso, **uma data precisa deve ser estipulada, e não atrelada a uma ordem do poder público que não tem prazo definido**, que é o fim da situação de calamidade pública do COVID-19.

Os dados coletados pelo IBGE deverão ser mantidos apenas durante o período que serão utilizados para fins de divulgação de cada relatório do PNADC. Logo, se os dados de uma divulgação mensal não forem necessários para a divulgação trimestral ou anual, deverão ser eliminados. A coleta desses dados, da forma como prevê a MP, não pode de forma alguma perdurar caso seja decretado o fim da calamidade pública.

VI - Medidas adicionais de proteção de dados

A MP 954 já traz, em seu escopo, alguns parâmetros de proteção de dados a serem seguidos pelo IBGE no tratamento dos dados pessoais, como data para eliminação, vedação de compartilhamento e obrigatoriedade de seguir a finalidade designada, que, como dito acima, deve ser melhor descrita.

A medida também prevê a realização de um relatório de impacto de proteção, o que é louvável, e faz expressa referência à LGPD, em seu art. 3º, §2º, o que demonstra intenção em seguir as diretrizes da lei.

Apesar disso, salvaguardas adicionais devem ser adotadas, conforme será descrito a seguir.

a. Nomeação de encarregado de proteção de dados

A primeira delas é a designação de um encarregado de proteção de dados. O ocupante do cargo será indicado pelo controlador e operador para atuar como canal de comunicação entre o controlador, os titulares dos dados e uma auditoria

independente, na ausência da Autoridade Nacional de Proteção de Dados (ANPD).

Seu objetivo é garantir que os tratamentos de dados realizados por uma organização sejam feitos de acordo com princípios e regras de proteção de dados adequadas. No Brasil, apesar de ainda não estar em pleno vigor, a LGPD traz esses parâmetros, e dispõe sobre a figura do encarregado em seus art. 5º, VIII, art. 23, III, e art. 41.

O encarregado deve ser um especialista na área e estar posicionado na estrutura hierárquica da organização de modo a poder trabalhar com independência e não receber ordens sobre seu trabalho.²¹ Além disso, o art. 23, III, da LGPD, estipula que o órgão público deve indicar um encarregado como condição para tratar dados pessoais.

b. Realização de auditoria externa

Sem a ANPD em funcionamento para trabalhar na fiscalização de atividades de tratamento de dados, é essencial que haja um colegiado de especialistas em proteção de dados sem vínculos com o IBGE para auditar a atuação do órgão. A auditoria externa também pode ser responsável por receber solicitações de titulares de dados que envolvam, por exemplo, seu direito de acesso ou de correção de suas informações.

A presença de uma auditoria será fundamental para aumentar o nível de confiança da sociedade²² tanto em relação ao IBGE quanto às próprias empresas de telecomunicações. Além disso, irá assegurar **maior transparência** a respeito de como os dados estão sendo tratados. Especialmente em uma época em que desinformação tem sido disseminada indiscriminadamente, ter informações precisas de como o governo usa nossos dados é essencial.

²¹ European Data Protection Supervisor. **Data Protection Officer (DPO)**. Disponível em https://edps.europa.eu/data-protection/data-protection/reference-library/data-protection-officer-dpo_en. Acesso em 22 abr. 2020.

²² European Data Portal. **Analytical Report 12: Business-to-Government Data Sharing**, p. 21. Mar. 2019. Disponível em https://www.europeandataportal.eu/sites/default/files/analytical_report_12_business_government_data_sharing.pdf. Acesso em 22 abr. 2020.

c. Estabelecer parâmetros de controle de acesso e de segurança dos dados

A MP 954 também deve prever que o IBGE aponte parâmetros de segurança da informação a serem seguidos durante a realização do tratamento de dados especificado. O órgão deve delimitar com clareza quem terá acesso aos dados, bem como quais as medidas de autenticação e de autorização necessárias para acessá-los. O IBGE também deve apontar quem será responsabilizado por eventuais abusos realizados com os dados ou vazamentos.²³

Mecanismos de segurança dos dados também devem ser previstos nas normas, especialmente pelo fato de o tratamento de dados, como explicitado acima, incluir dados sensíveis. O IBGE deve ser transparente com a forma como protege dados contra invasões e vazamentos, considerando o fato de pertencerem a uma categoria especial.

d. Anonimização e pseudoanonimização

Uma hipótese a ser analisada é a utilização de mecanismos de anonimização e pseudoanonimização durante o processo de realização das estatísticas.

O conceito de anonimização de dados se refere, nos termos da LGPD, à “utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo”. Ao serem anonimizados, dados deixam de ser pessoais, e não estão mais sujeitos às regras da LGPD.

A esse respeito, vale ressaltar que o processo de anonimização de dados pessoais nunca é perfeito. Conforme descrito por Finck e Pallas,²⁴ a anonimização deve ser entendida como um meio de **reduzir o risco** de identificação de titulares de dados. Embora nunca seja absoluta, a anonimização apenas ocorrerá se puder garantir que, em um determinado período, os recursos técnicos e financeiros

²³ *Idem.*

²⁴ FINCK, Michèle; PALLAS, Frank. **They who must not be identified—distinguishing personal from non-personal data under the GDPR**. International Data Privacy Law, Oxford, pp. 1-26, 10 mar. 2020. Disponível em: <https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf>. Acesso em: 9 abr. 2020.

então existentes tornariam o processo de re-identificação impraticável ou excessivamente dispendioso.

Ou seja, **a anonimização só existe se o risco de re-identificar um indivíduo for residual e irrelevante.** Baseado nessa abordagem, a agregação é apenas mais uma técnica (embora bastante relevante neste contexto) para garantir a anonimização dos dados. Nesse sentido, o emprego de métodos de agregação, como *k-anonymity*, *l-diversity* e *t-closeness* devem ser incentivados mesmo em compartilhamentos dentro do próprio órgão.

Por outro lado, para dados que não poderão ser anonimizados antes de serem eliminados ao fim da pesquisa, técnicas de **pseudoanonimização** podem ser aplicadas.

De acordo com o art. 13, §4º, da LGPD, pseudoanonimização é 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.

Trata-se de um método de tratamento protetivo para dados. Apesar de não perderem o caráter pessoal, a pseudoanonimização garante maiores obstáculos para re-identificar indivíduos. Exemplos de aplicação são o emprego de recursos de criptografia e do estabelecimento de *hashes*, por exemplo.

A adoção de tais técnicas pelo IBGE é recomendada considerando tanto o alto volume de dados coletados quanto sua natureza sensível. Caso mecanismos adequados para impedir sua re-identificação não sejam postos em prática, aumenta o risco de cruzamento dessa base de dados com outras a fim de descobrir a identidade de indivíduos afetados pela Medida Provisória.

Isso pode dar margem tanto para uma violação à finalidade previamente estabelecida de tratamento desses dados, como também a possibilidade de levar indivíduos a serem discriminados com base em sua raça, orientação sexual ou sobre suas informações de saúde.

Nesse sentido, o risco de se ter uma base de dados que permita uma violação em massa da privacidade das pessoas é muito alto para que não se adote técnicas preventivas, em respeito ao princípio da segurança e da prevenção

previstos no art. 6, incisos VII e VIII da Lei Geral de Proteção de Dados.

e. Adoção de código de conduta

Na ausência da LGPD, um código de conduta deveria ser adotado pelo IBGE, de modo a tornar transparente o fato de que tem seguido princípios claros de proteção de dados. Um exemplo de código a ser seguido é o descrito pela Comissão Europeia para transferências de dados entre organizações, que inclui, entre outros princípios, os de proporcionalidade, finalidade e transparência.²⁵

f. Garantia do exercício dos direitos do titular de dados

Além dos princípios estabelecidos no art. 6º da LGPD, o IBGE deve garantir o pleno exercício dos direitos do titular, previstos no art. 18 da mesma lei, como o direito de **livre acesso** aos dados e de **retificação**.

Outro direito importante de ser garantido é o de **revogação** do consentimento, caso a entrevista seja realizada de forma voluntária. O titular de dados entrevistado tem o direito de manifestar sua desistência do interesse em participar da pesquisa.

Para tanto, conforme descrito acima, a criação de um canal aberto de comunicação, que pode ser exercido pela auditoria externa, para lidar com o caso específico das pesquisas realizadas durante COVID-19 deveria ser prevista no procedimento de tratamento de dados. Cumprir com o princípio da **transparência**, nesse sentido, é fundamental.

Pelo canal, o titular de dados poderia encaminhar suas solicitações, que deveriam ser cumpridas em tempo hábil, preferencialmente dentro de 15 dias, de modo a garantir ao indivíduo controle sobre suas informações. Esse prazo está de acordo com o estipulado no art. 19, II, da LGPD.

²⁵ European Commission. **Guidance on sharing private sector data in the European data economy**. Disponível em <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy>. Acesso em 22 abr. 2020.

VI - Conclusão

A despeito da necessidade de produção de dados estatísticos oficiais no Brasil, a Medida Provisória 954/2020 fere os princípios da necessidade, adequação e finalidade previstos no art. 6º da Lei Geral de Proteção de Dados. Isso porque falha em não especificar para qual pesquisa que os dados serão tratados, bem como por coletar um número excessivo de dados para os fins a que se destinariam.

A Medida Provisória também peca ao não prever salvaguardas adequadas para garantir que a utilização de tais dados seja feita de forma correta, dada a falta de indicação de auditoria externa e a ausência de parâmetros de segurança adequados para gerir os dados. Além disso, não há qualquer referência a mecanismos de anonimização de dados ou de adesão a código de conduta que garantam o direito à proteção de dados dos indivíduos.

Nesse sentido, as seguintes falhas, apontadas ao longo da Nota Técnica, devem ser superadas:

1. especificar a quais estudos estatísticos a MP se refere;
2. coletar com as operadoras telefônicas apenas os telefones dos clientes e mediante amostragem aleatória;
3. durante a entrevista, coletar dados de forma voluntária, com base no consentimento, de forma específica e destacada para a finalidade específica;
4. vincular a data de eliminação dos dados ao fim do estudo estatístico e não ao fim da pandemia;
5. nomear encarregado de proteção de dados;
6. realização de auditoria externa;
7. estabelecer parâmetros de controle de acesso e de segurança dos dados;
8. implementar técnicas de anonimização e pseudonimização;
9. adotar códigos de conduta; e
10. garantir o exercício dos direitos dos titulares de dados.

Pelo exposto, o LAPIN se posiciona pela correção das falhas apresentadas, de modo a garantir uma redação mais adequada aos parâmetros previstos pela LGPD e por normas de proteção de dados ao redor do globo.