UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE DIREITO
DEPARTAMENTO DE DIREITO PRIVADO E PROCESSO CIVIL

Eduarda Beutinger Paiva

A REVERSIBILIDADE DO PROCESSO DE ANONIMIZAÇÃO E AS SUAS REPERCUSSÕES NO REGIME DE PROTEÇÃO DE DADOS PESSOAIS

Porto Alegre
2021

Eduarda Beutinger Paiva

A REVERSIBILIDADE DO PROCESSO DE ANONIMIZAÇÃO E AS SUAS REPERCUSSÕES NO REGIME DE PROTEÇÃO DE DADOS PESSOAIS

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do grau de Bacharel em Ciências Jurídicas e Sociais na Universidade Federal do Rio Grande do Sul.

Orientador: Professor Doutor Fabiano Menke.

Porto Alegre

2021

Eduarda Beutinger Paiva

A REVERSIBILIDADE DO PROCESSO DE ANONIMIZAÇÃO E AS SUAS REPERCUSSÕES NO REGIME DE PROTEÇÃO DE DADOS PESSOAIS

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do grau de Bacharel em Ciências Jurídicas e Sociais na Universidade Federal do Rio Grande do Sul.

Orientador: Professor Doutor Fabiano Menke.

Aprovado em 24 de novembro de 2021.

BANCA EXAMINADORA:

_____

Prof. Dr. Fabiano Menke

Orientador

_____

Prof. Dr. Luis Renato Ferreira da Silva

_____

Prof. Dr. Gerson Luiz Carlos Branco

AGRADECIMENTOS

Resiliência é a palavra que melhor define o ano de 2021. Em meio à realidade caótica de uma pandemia, em que a incerteza tomava conta dos nossos dias, todos tivemos que nos adaptar e nos reinventar em diversos aspectos. Para mim não foi diferente. Desde o início do curso, planejei apresentar o meu Trabalho de Conclusão de Curso no oitavo semestre. Não aconteceu. Não foi por falta de vontade ou organização, mas pelo vazio que aquele período me causava. As notícias relatando o cotidiano nos hospitais, o número de mortes que crescia, as notas de falecimento que recebia diariamente no e-mail do estágio, a saudade dos meus parentes, amigos, colegas e professores que não encontrava há meses... A inaptidão de enxergar uma luz no fim do túnel fazia com que o processo de escrita fosse quase que impraticável. Por reiteradas vezes fiquei com o documento do *Word* aberto e o cursor piscando, sem nada digitar. Essa situação apenas foi revertida graças ao apoio que recebi, inclusive, de pessoas que sequer sabem que me ajudaram. Aos médicos que incansavelmente lutaram contra essa doença, aos pesquisadores que tanto se dedicaram para produzir a vacina, aos professores que do dia para a noite tiveram que se adaptar à realidade de aulas virtuais, aos trabalhadores que não tiveram outra opção senão trabalhar em meio à pandemia, àqueles que ficaram meses no combate contra a doença e venceram, a todas essas pessoas que eu nem mesmo conheço: OBRIGADA. Obrigada por me motivarem e por permitirem que eu percebesse que, por mais difícil que seja, é possível nós mesmos acendermos a nossa própria luz no fim do túnel.

Do agradecimento geral passo para o mais pessoal. Aos meus pais, obrigada pelo apoio e por sempre acreditarem na minha capacidade. Ambos são grandes exemplos para mim. Ao meu pai, Ely Paiva, por sempre ser minha grande referência no mundo acadêmico. Minha admiração por ti é imensurável. À minha mãe, Rejane Paiva, por ser o meu porto seguro em todos os momentos da minha vida. À minha irmã, Daniela, que atualmente inicia sua trilha no mundo jurídico e que em muito me orgulha. Aos meus avós, que, mesmo de longe, se fazem presente e me encorajam a nunca desistir dos meus sonhos. À minha tia, Luciana Paiva, primeira mulher a assumir o cargo de Direção da ESEFID/UFRGS, que foi e é fundamental em todo o meu desenvolvimento acadêmico e que é uma das figuras femininas que mais me inspiram. Ao meu amigo, colega de curso e parceiro de vida, Pedro, pelo apoio, companheirismo e por sempre estar ao meu lado seja qual for o momento. Aos meus amigos do colégio e da faculdade por sempre me motivarem e torcerem por mim. Aos colegas de

grupo de estudos, especialmente, Anita, Bibiana e Rafael, pela ajuda em distintas fases deste trabalho. Aos meus colegas de trabalho pela compreensão e motivação neste momento.

Por fim, gostaria de agradecer ao meu grande mestre que, durante toda a faculdade, foi e é minha referência na Academia de Direito: Professor Dr. Fabiano Menke. Obrigada pela dedicação ímpar ao ensino, à pesquisa e aos seus alunos. A semente que nos é plantada em suas aulas e Grupos de Estudos é cultivada para o resto de nossas vidas. Desde o meu primeiro SIC, em 2018, até o TCC, escolhi o senhor como orientador e sinto que minha escolha não poderia ser melhor. Minha vontade em continuar a carreira acadêmica e meus estudos na Língua Alemã é em muito motivada pela sua trajetória. Muito obrigada.

RESUMO

O presente estudo tem como objetivo central analisar a maneira pela qual a Lei Geral de Proteção de Dados Pessoais – LGPD realiza a diferenciação entre os conceitos de dado pessoal e de dado anonimizado, tendo em vista a reversibilidade do processo de anonimização, e como isso repercute no escopo de aplicação do regime de proteção de dados pessoais. Para tanto, dividiu-se o trabalho em três partes. Em primeiro momento, delimita-se o conceito de dado pessoal, realizando uma análise individualizada de cada um dos elementos que compõem tal definição. Introduz-se, então, a ideia de dualidade mutualmente excludente entre os conceitos de dado pessoal e de dado anonimizado, além de diferenciar este dos conceitos de dado pseudonimizado e de dado criptografado. Na segunda parte, aborda-se sobre a reversibilidade do processo de anonimização e os seus efeitos no plano legal, essencialmente, no tocante à forma pela qual a legislação brasileira e o Regulamento Europeu de Proteção de Dados – GDPR lidam com o risco inerente de um dado anonimizado se transmudar em um dado pessoal. Cita-se estudos relevantes da área da ciência da computação que comprovam as fragilidades das técnicas de anonimização de dados que, em muitos casos, permitem a reidentificação da pessoa natural à qual os dados se referem. Na terceira e última parte, apresenta-se o critério adotado pela LGPD e pelo GDPR para diferenciar dados pessoais de dados anonimizados, mesmo sendo reconhecido o risco da reversão do processo de anonimização: a razoabilidade. São indicados os dois eixos essenciais de análise, quais sejam o eixo subjetivo e o eixo objetivo, cada um com parâmetros próprios previstos em Lei. Por fim, dispõe-se sobre como a anonimização de dados, ainda que não reduzindo a zero a possibilidade de reidentificação do titular, pode ser uma grande aliada na mitigação dos riscos advindos do tratamento de dados pessoais.

Palavras-chave: Proteção de dados pessoais. Anonimização. Reidentificação. Critério da razoabilidade. Mitigação de riscos.

ABSTRACT

This study aims to analyze how the concepts of personal data and anonymized data are distinguished according to the Brazilian Data Protection Regulation - LGPD, taking into account the reversibility of the anonymization process. The study seeks also to demonstrates how anonimization affects the scope of application of the personal data protection regime. The monography adresses the discussion in three topics. Firstly, the concept of personal data is delimited, carrying out an individual analysis of each of the elements that compose such definition. The idea of a mutually exclusive duality between the concepts of personal data and anonymized data is then introduced, in addition to differentiating the latter from the concepts of pseudonymized and encrypted data. The second part deals with the reversibility of the anonymization process and its effects on the legal plane, essentially, with respect to the manner in which Brazilian legislation and the UE General Data Protection Regulation - GDPR deal with the inherent risk of anonymized data transmuting into personal data. Relevant computer science studies are cited, presenting the weaknesses of data anonymization techniques that in many cases allow the re-identification of the natural person to whom the data relates to. Finally, the criteria adopted by the LGPD and the GDPR to differentiate personal data from anonymized data, even though the risk of reversing the anonymization process is recognized: reasonableness. The two essential axes of analysis are indicated, namely the subjective axis and the objective axis, each one with its own parameters provided by Law. It is discussed how data anonymization, even if it does not reduce to zero the possibility of re-identification of the data subject, can mitigate the arising risks from the personal data processing.

Keywords: Personal data protection. Data anonymization. Re-identification. Reasonable likelihood. Risk mitigation.

SUMÁRIO

1. INTRODUÇÃO

Em 18 de setembro de 2020, entrou em vigor a Lei Geral de Proteção de Dados Pessoais – LGPD (Lei nº 13.709, de 14 de agosto de 2018), trazendo consigo, para a mais variada gama de setores, desafios quanto à sua implementação prática. Desde a sua *vacatio legis*, a adequação à LGPD tornou-se objeto de cursos, palestras e consultorias que visam capacitar profissionais, tanto no âmbito público quanto no privado, para conhecer os requisitos previstos na Lei e melhor atendê-los.

Conforme prevê o artigo 5º, inciso I, da LGPD, dado pessoal é "informação relacionada a pessoa natural identificada ou identificável". Tal conceito consiste no núcleo duro da Lei e delimita o seu escopo de aplicação, já que essa se destina à proteção dos direitos da pessoa natural titular de dados pessoais. Diante de uma realidade crescentemente digitalizada, são poucas as empresas ou mesmo órgãos do setor público que não realizam o tratamento de dados pessoais, sendo, portanto, imperativo que toda essa diversidade de atividades (não apenas aquelas ligadas à tecnologia, mas também à saúde, setor financeiro, segurança, etc.) entre em conformidade com o regime de proteção de dados pessoais. Para tanto, muitos destes agentes passaram a recorrer a meios de mitigação dos riscos provenientes do tratamento de dados.

Uma medida que está sendo amplamente utilizada é a implementação de PETs (*Privacy Enhacing Technologies*). Trata-se de um termo "guarda-chuva" que abrange todas as tecnologias que facilitam a privacidade e a segurança da informação. Elas se encontram no âmbito da metodologia do *Privacy by Design*, que se traduz na ideia de arquitetar produtos e serviços que sejam construídos, desde o princípio, com a adoção de tecnologias que promovam o controle e a proteção dos dados pessoais[1]. A proposta central das PETs é preservar o valor do tratamento dos dados pessoais e, consequentemente, o interesse que tal atividade produz nos agentes de tratamento, sem comprometer os direitos de personalidade do titular.

As técnicas de anonimização de dados constituem PETs que, apesar de terem potencial para se tornarem grandes aliadas na garantia de maior grau de *compliance* com a Lei, despertam dúvidas no tocante à sua implementação prática. Isso porque dados anonimizados, ou seja, dados pessoais que passaram por processo de retirada total ou parcial dos vínculos com o titular, visando impedir ou, pelo menos, dificultar a sua identificação, estão fora do escopo de aplicação da LGPD, não sendo, pois, objeto do seu regime de proteção. Assim,

[1] BIONI, Bruno. **Proteção de dados pessoais: a função e os limites do consentimento.** 1. ed. Rio de

muitos agentes de tratamento presumem que, por meio da adoção de técnicas de anonimização dos dados que dispõe, estariam se esquivando dos regulamentos da Lei e, consequentemente, de sanções em razão de possíveis vazamentos. O problema advém, todavia, do fato de o processo de anonimização poder ser revertido, o que significa dizer que dados anonimizados podem recuperar a possibilidade de serem associados a uma pessoa, seja pela combinação com outros dados disponibilizados ou pela própria evolução das (contra)tecnologias disponíveis, de modo a se tornarem, novamente, dados pessoais.

Tendo em vista este cenário, o presente trabalho tem como objetivo responder à seguinte questão: considerando que o conceito de dado pessoal, previsto no artigo 5º, inciso I, da LGPD, abarca a possibilidade de o titular dos dados ser meramente identificável, como pode ser feita a diferenciação entre esta categoria e a dos dados anonimizados, uma vez que o processo de anonimização pode ser revertido, tornando o titular, da mesma forma, identificável?

Consoante se depreende do seu artigo 12, *caput*[2], a LGPD reconhece a existência do risco residual de o processo de anonimização ser revertido, o que torna necessária a adoção de elementos adicionais que permitam a sua distinção do conceito de dado pessoal. Nota-se, no citado dispositivo, a presença de dois elementos principais que permitem tal diferenciação: a utilização de meios próprios (eixo subjetivo) ou o emprego de esforços razoáveis (eixo objetivo). Conforme disposto no referido artigo, dados anonimizados passam a ser considerados dados pessoais se o processo de anonimização pelo qual passaram for revertido mediante o uso de meios próprios, ou se ele puder ser revertido com esforços razoáveis.

Utilizando-se do método de abordagem dedutivo, o estudo tem como objetivo analisar ambos os critérios, verificando de que forma estes parâmetros são delimitados quando da implementação de técnicas de anonimização e se a sua não observação é suficiente para afastar o regime de proteção de dados pessoais. Quanto ao método de procedimento, foi realizada a pesquisa bibliográfica e documental, por meio da análise de doutrina nacional e estrangeira, especialmente, da União Europeia e dos Estados Unidos da América. Ainda, utilizar-se-á o método comparativo funcional visando a comparação da forma pela qual a reversibilidade do processo de anonimização é encarada pela legislação nacional àquela adotada no Regulamento da União Europeia (GDPR – *General Data Protection Regulation*).

---

[2] Art. 12. Os dados anonimizados não serão considerados dados pessoais para os fins desta Lei, salvo quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

A exposição do estudo será dividida em três partes. No primeiro capítulo, delimita-se o conceito de dado pessoal, realizando uma análise individualizada de cada um dos elementos que compõem tal definição. Introduz-se, então, a ideia de dualidade mutualmente excludente entre os conceitos de dado pessoal e de dado anonimizado, além de diferenciar este dos conceitos de dado pseudonimizado e de dado criptografado. No segundo capítulo, aborda-se sobre a reversibilidade do processo de anonimização e os seus efeitos no plano legal, essencialmente, no tocante à forma pela qual a legislação brasileira e o Regulamento Europeu de Proteção de Dados – GDPR lidam com o risco inerente de um dado anonimizado se transmudar em um dado pessoal. Cita-se estudos relevantes da área da ciência da computação que comprovam as fragilidades das técnicas de anonimização de dados que, em muitos casos, permitem a reidentificação da pessoa natural a qual os dados se referem. No terceiro e último capítulo, apresenta-se o critério adotado pela LGPD e pelo GDPR para diferenciar dados pessoais de dados anonimizados, mesmo sendo reconhecido o risco da reversão do processo de anonimização: a razoabilidade. São indicados os dois eixos essenciais de análise, quais sejam o eixo subjetivo e o eixo objetivo, cada um com parâmetros próprios previstos em Lei. Por fim, dispõe-se sobre como a anonimização de dados, ainda que não reduzindo a zero a possibilidade de reidentificação do titular, pode ser uma grande aliada na mitigação dos riscos advindos do tratamento de dados pessoais.

2. O REGIME LEGAL DE PROTEÇÃO DE DADOS PESSOAIS

A preocupação com relação ao comprometimento de direitos e garantias de indivíduos por conta do tratamento de informações ligadas a estes remonta à década de 1960, momento em que, nos Estados Unidos da América, se discutia a necessidade de limitação do poder do Estado na vida privada dos cidadãos. O termo "*privacy*" assumia uma conotação individualista, sendo utilizado para designar o controle do indivíduo sobre as suas informações pessoais, o que tinha, precipuamente, a função de garantir a sua autonomia e liberdade em face das arbitrariedades do Estado[3]. Neste período, a ideia relativa à "proteção de dados pessoais" detinha sentido essencialmente prático, referindo-se às medidas técnicas adotadas para proteger os dados pessoais, e não, propriamente, a um direito autônomo[4]. A proteção dos dados pessoais estava, portanto, relacionada à segurança da informação, sendo um instrumento para preservar a vida privada dos cidadãos frente ao poder estatal.

A partir de 1970, com a aprovação da *Hessiches Datenschutzgesetsz – HDSG* pelo Parlamento do *Land* Alemão de Hesse, novos rumos passaram a ser traçados. A expressão "proteção de dados" (*Datenschutz*) foi pela primeira vez utilizada e um modelo normativo autônomo destinado à proteção de dados pessoais foi pioneiramente reconhecido. Seguindo os embates que a nomenclatura adotada pela referida lei despertou na Alemanha, Ulrich Seidel, em sua obra "*Persönlichkeitsrechtliche Probleme der elektronischen Speicherung privater Daten*" de 1972, defendeu a necessidade de diferenciar os termos "*Datensicherung*" e *"Datenschutz*" que correspondem, respectivamente, à proteção técnica de dados e à proteção jurídica concedida aos titulares de dados[5]. A terminologia "*Datenschutz*" consolidou-se na Alemanha e na Europa[6] e, gradualmente, o direito à proteção de dados pessoais foi adquirindo moldes de um direito fundamental autônomo que ia além da segurança da informação, da privacidade e do sigilo[7].

A necessidade de regular autonomamente a atividade de tratamento de dados pessoais e de, consequentemente, reconhecer um direito autônomo à proteção de dados pessoais surgiu com a evolução das tecnologias da informação, que permitiram um processamento massivo e

---

[3] FUSTER, Gloria González. **The Emergence of Personal Data Protection as a Fundamental Right of the EU**. Bruxelas: Springer Science & Business, 2014. p. 23.
[4] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 31.
[5] *Ibidem*.
[6] *Ibidem*. p. 32.
[7] DONEDA, Danilo. Panorama histórico da proteção de dados pessoais. In: BIONI, Bruno et al. (coord.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. p. 8.

contínuo de dados[8]. Desde o advento da disciplina, os riscos envolvendo a atividade aumentaram quantitativamente e qualitativamente por conta do volume de dados disponibilizados e do aperfeiçoamento das técnicas de tratamento. O tratamento manual deu lugar ao tratamento automatizado[9], e as fichas manuscritas deram lugar ao *Big Data*[10]. Além disso, passou-se a lidar com fontes de dados que sequer existiam há uma geração atrás, tais como redes sociais e *smartphones*; e com métodos de análise de dados que estão evoluindo para se adequar às novas fontes (*cloud computing, machine learning*, ...)[11].

A datificação, ou seja, a transformação de toda e qualquer informação em dados que podem ser lidos, manipulados ou combinados por inteligência artificial[12], trouxe mudanças fundamentais à vida social, à economia e a diversas outras esferas da realidade material. O Direito, como garantidor da convivência ordenada entre os membros que compõe a sociedade e as suas instâncias[13], não poderia se manter inerte frente a tal fenômeno. Ao redor do globo, diversas legislações de proteção de dados pessoais passaram a ser elaboradas, alcançando a marca atual de 128 países com diplomas normativos que tratam especificamente sobre a matéria[14]. Na maior parte destes diplomas, a disciplina jurídica da proteção de dados pessoais está embasada na imprescindibilidade do mercado de dados para o desenvolvimento da sociedade. Logo, pode-se dizer que o regime legal de proteção de dados pessoais possui dois

---

[8] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 53.

[9] O tratamento automatizado de dados baseia-se na ideia de decisão por meios automatizados, sem qualquer envolvimento humano. Essas decisões podem ser baseadas em dados factuais, em perfis criados digitalmente (*profiling*) ou em combinações de dados. Vide: UK INFORMATION COMMISSIONER'S OFFICE – ICO. **What is automated individual decision-making and profiling?**. Disponível em: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/>. Acesso em 20 out. 2021: "Automated decision-making is the process of making a decision by automated means without any human involvement. These decisions can be based on factual data, as well as on digitally created profiles or inferred data.".

[10] Definir *Big Data* é uma tarefa complexa, pois não há uma delimitação concreta deste fenômeno. De acordo com Bart van der Sloot, Dennis Broeders e Erik Schrijvers, a definição de *Big Data* dividi-se em três aspectos: **(i)** coleta de dados – o *Big Data* diz respeito à coleta de grandes quantidades de dados, a partir de fontes variadas e, muitas vezes, não estruturadas; **(ii)** análise de dados – o Big Data relaciona-se, também, à velocidade das análises e do uso de certos mecanismos como algoritmos, *learning machine* e correlações estatísticas; **(iii)** uso de dados - os resultados são muitas vezes de natureza preditiva e são formulados em nível geral ou de determinado grupo. Vide: VAN DER SLOOT, Bart; BROEDERS, Dennis; SCHRIJVERS, Erik (org.). **Exploring the boundaries of Big Data**. Amsterdam: WRR, 2016. p. 11.

[11] FLORIDI, Luciano; TAYLOR, Linnet; VAN DER SLOOT, Bart. Introduction: A New Perspective on Privacy. In: FLORIDI, Luciano; TAYLOR, Linnet; VAN DER SLOOT, Bart (coord.). **Group Privacy**: New Challenges of Data Technologies. Dordrecht: Springer International Publishing, 2017. *E-book*. p. 3.

[12] *Ibidem*.

[13] REALE, Miguel. **Lições preliminares de direito**. 27. ed. São Paulo: Saraiva, 2002. p. 15.

[14] Consoante pesquisa realizada no âmbito da *United Nations Conference on Trade and Development* – UNCTAD, 128 dos 194 países verificados no estudo já implementaram leis para garantir a proteção de dados e da privacidade. UNCTAD. **Data Protection And Privacy Legislation Worldwide**. 2020. Disponível em: <https://unctad.org/page/data-protection-and-privacy-legislation-worldwide>. Acesso em 09 set. 2021.

propósitos principais: (i) defender os direitos e liberdades fundamentais das pessoas naturais titulares de dados pessoais; (ii) promover a livre circulação dos dados pessoais[15]. O objetivo é atender, em harmonia, os interesses dos titulares e os interesses das entidades públicas e privadas que dependem, para o exercício das suas atividades, do tratamento de dados.

Já no ano de 1999, Colin Bennett e Rebecca Grant, em sua obra "*Visions Of Privacy: Policy Choices For The Digital Age*", apontavam a existência de uma tendência à convergência internacional entre as políticas destinadas à regulação do tratamento de dados e da privacidade[16]. De acordo com os autores, apesar de haver diferenças nas regulamentações, especialmente, no tocante a questões relativas à nomenclatura e à implementação, aquilo que é essencial encontra-se sempre presente[17]. Nesse sentido, observa-se a existência de um conceito fundamental em todos os diplomas, que pode ser chamado de "núcleo duro" do sistema de proteção de dados: o conceito de dado pessoal.

2.1 O CONCEITO DE DADO PESSOAL

O regime de proteção de dados pessoais tem como razão de ser o conceito de dado pessoal. Este elemento fundamental define o escopo material de aplicação da proteção concedida pela LGPD e por todos os demais diplomas relativos à matéria, tais como o Regulamento Geral de Proteção de Dados Pessoais da União Europeia (*General Data Protection Regulation – GDPR*) e a Convenção 108 do Conselho da Europa[18]. Todas as citadas regulações adotam um conceito padrão de dado pessoal, qual seja, toda informação relativa a pessoa natural identificada ou identificável. Consoante afirma Bruno Bioni, "compreender se um dado pode ser adjetivado como pessoal é, antes de tudo, um exercício de interpretação detido sobre cada palavra utilizada para prescrever a sua conceituação"[19]. Assim, adotando o método utilizado pelo Grupo de Trabalho do Artigo 29 (*Article 29 Data Protection Working Party* – WP29)[20], no Parecer 4/2007 sobre o conceito de dados pessoais[21],

---

[15] FUSTER, Gloria González. **The Emergence of Personal Data Protection as a Fundamental Right of the EU**. Bruxelas: Springer Science & Business, 2014, p. 21-54. *E-book*.
[16] BENNET, Colin J.; GRANT, Rebecca. **Visions Of Privacy: Policy Choices For The Digital Age**. Toronto: University of Toronto Press - Scholarly Publishing Division, 1999. p. 5-8.
[17] *Ibidem*.
[18] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e criptografia: tecnologias criptográficas entre anonimização e pseudonimização de dados. In: DONEDA, Danilo; MACHADO, Diego (coord.) **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019, p. RB-8.2. *E-book*.
[19] BIONI, Bruno. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, 2015. p. 17.
[20] O Grupo de Trabalho do Artigo 29 (*Article 29 Working Party* – WP29) foi um comitê criado pela Diretiva 95/46/CE que tinha como papel lidar com questões relativas à interpretação das normas da Diretiva, resolvendo,

far-se-á a análise individual de todos os pilares que compõe a definição[22], quais sejam: (i) informação; (ii) relacionada; (iii) pessoa natural; (iv) identificada ou identificável.

2.1.1 Informação

De início, cumpre estabelecer as diferenças entre os termos "dado" e "informação", que, em inúmeras circunstâncias, são utilizados como sinônimos, apesar de não o serem. Ambos representam determinado aspecto da realidade[23], e, por vezes, o seu conteúdo se sobrepõe, razão pela qual é notável certa promiscuidade na sua utilização[24]. Salienta-se, todavia, que o dado corresponde a um estágio anterior à informação. Em outras palavras, o dado é uma espécie de "pré-informação", passando a adquirir o *status* de informação após o processo de interpretação[25]. A informação, por sua vez, está a uma etapa antes do conhecimento. Assim, para facilitar a visualização e o entendimento, pode-se imaginar uma figura piramidal, em que o grau de abstração da base ao topo é decrescente. Na base da pirâmide, encontra-se o termo "dado", e no seu cume o termo "conhecimento". Entre tais termos fica localizada a "informação".

A doutrina e a lei, recorrentemente, fazem o uso dos termos "dado" e "informação" de maneira indistinta, o que é perceptível, inclusive, nos conceitos de "dado pessoal" trazido pela LGPD[26] e de "informação pessoal" trazido pela Lei de Acesso à Informação[27] que são,

---

por exemplo, ambiguidades constantes no texto e esclarecendo a aplicação de determinados conceitos. Apesar de suas atividades terem sido encerradas em maio de 2018, quando, com a vigência do GDPR, foi substituído pelo Conselho Europeu para a Proteção de Dados (*European Data Protection Board* - EDPB), os pareceres elaborados continuam válidos e aplicáveis sob o novo Regulamento. Vide: EUROPEAN COMMISSION. **The Article 29 Working Party ceased to exist es of 25 May 2018**. Disponível em: <https://ec.europa.eu/newsroom/article29/items/629492/en>. Acesso em 07 set. 2021; e DOLIN, Ron A. Search Query Privacy: The Problem of Anonymization. **Hastings Science and Technology Law Journal**, v. 2, n. 2, p. 137-182, mai. 2010. p. 140.

[21] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

[22] Tendo em vista que o objeto da presente monografia é a Lei brasileira, o conceito a ser analisado no presente subcapítulo é o constante no inciso I do artigo 5º da LGPD. "dado pessoal: informação relacionada a pessoa natural identificada ou identificável". BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[23] DONEDA, Danilo. A proteção dos dados pessoais como um direito fundamental. **Espaço Jurídico**, v. 12, n. 2, p. 91-108, jul./dez. 2011. p. 94.

[24] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019. p. 136.

[25] *Ibidem*.

[26] Art. 5º Para os fins desta Lei, considera-se: I - dado pessoal: informação relacionada a pessoa natural identificada ou identificável; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

praticamente, idênticos[28]. Para a efetivação da tutela jurídico-constitucional, a distinção entre dados e informações torna-se irrelevante. O que importa é a configuração dos requisitos legais referidos, e não a forma mediante a qual se corporifica o aspecto da realidade representado[29]. Portanto, no que tange à aplicação da LGPD, a relevância está no fato de o dado ou informação possuir vínculo a uma pessoa, revelando algum aspecto objetivo desta[30].

Conforme proposto por Pierre Catala, a informação, a depender do seu conteúdo, pode ser classificada em quatro modalidades: (i) as informações relativas às pessoas e seus patrimônios; (ii) as opiniões subjetivas das pessoas; (iii) as obras do espírito; (iv) as descrições de fenômenos e coisas[31]. O primeiro grupo corresponde àquelas informações caracterizadas como pessoais, isso é, que abrangem todos os aspectos relativos a determinado indivíduo, sejam eles familiares ou sociais, privados ou públicos, físicos ou mentais[32]. Para que sejam consideradas como pessoais, as informações devem ter como objeto a própria pessoa, estabelecendo com esta um vínculo objetivo. É esse vínculo que distingue a categoria das demais existentes, porquanto, há informações que, apesar de fazerem referência a determinada pessoa, não são consideradas informações pessoais. A título exemplificativo, tem-se que a produção intelectual de uma pessoa, *per se*, não é considerada informação pessoal, a autoria, por sua vez, o é[33].

Para efeitos de aplicação da LGPD, toda informação pessoal é relevante e merecedora de proteção jurídica, ainda que isoladamente aparente ser insignificante. Esta máxima vem sendo consolidada desde a histórica decisão da Lei do Censo (*Volkszählungsurteil*), em julgamento realizado pelo Tribunal Constitucional Federal Alemão em 15.12.1983, oportunidade em que restou reconhecido que "não existem mais dados "insignificantes" no contexto do processamento eletrônico de dados"[34]. Com o advento de sistemas de informação

---

[27] Art. 4º Para os efeitos desta Lei, considera-se: [...] IV - informação pessoal: aquela relacionada à pessoa natural identificada ou identificável; [...]. BRASIL. **Lei nº 12.527, de 18 de novembro de 2011.**. Lei de Acesso à Informação (LAI). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em 01 set. 2021.

[28] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019. p. 136.

[29] SARLET, Ingo Wolfgang. Fundamentos constitucionais: o direito fundamental à proteção de dados. In: BIONI, Bruno et al. (coord.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. p. 39-40.

[30] DONEDA, Danilo. A proteção dos dados pessoais como um direito fundamental. **Espaço Jurídico**, v. 12, n. 2, p. 91-108, jul./dez. 2011. p. 94.

[31] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019. p. 139.

[32] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 108.

[33] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019. p. 139

[34] DEUTSCHLAND. BVerfGE 65, 1. Bundesverfassungsgericht, Karlsruhe, 1983. Disponível em: <https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/1983/12/rs19831215_1

integrados, a combinação de dados tornou-se algo facilmente alcançável, o que permite que dados aparentemente "inofensivos" possam ser processados para as mais diversas finalidades, adquirindo novo valor e, consequentemente, novo grau de risco ao seu titular. Logo, não se pode condicionar a aplicação do regime de proteção de dados pessoais a determinados tipos de informação pessoal. Desde que revele algum aspecto da personalidade de um indivíduo, a informação é considerada pessoal e, portanto, aplica-se as normas de proteção estabelecidas na LGPD.

No mesmo sentido, não se admite a diferenciação entre informações privadas e informações públicas no tocante à aplicação da Lei. Historicamente, em razão de o nascimento da disciplina de proteção de dados estar diretamente conectado com o direito à privacidade, sustentava-se existir uma extensão da tutela de tal direito à proteção de dados pessoais, o que acabava limitando o seu alcance àqueles dados tidos como sigilosos ou privados[35]. Todavia, este entendimento não mais abrange a complexidade atual do fenômeno do tratamento de dados, de modo que a tutela garantida ao titular de dados se tornaria frágil e superficial caso fosse concedida apenas a dados de caráter sigiloso. O número de fontes e destinatários de dados multiplicou-se com a evolução das tecnologias da informação, o que, inclusive, dificultou a diferenciação entre dados públicos e privados, tornando descabida esta classificação binária. Tal "ampliação" da proteção destinada a dados pessoais já foi reconhecida pelo Supremo Tribunal Federal, quando do julgamento das ADIs 6.387, 6.388, 6.389, 6.390 e 6.393, em maio de 2020. Na oportunidade, a ministra Cármen Lúcia, em seu voto, sintetiza, metaforicamente, a situação atual: "[...] não mais estamos em tempos em que é possível a publicização de dados em listas telefônicas sem a constituição de riscos ao cidadão; hoje, qualquer tipo de dado, ainda que publicizado, pode ser utilizado contra os interesses daquele que o titula"[36-37].

O conceito de informação pessoal abrange, portanto, todas as esferas e modalidades de aspectos relativos ao indivíduo, desde suas características físicas e seu nome até

---

bvr020983.html>. Acesso em 20 set. 2021.

[35] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2 ed. São Paulo: Thomson Reuters Brasil, 2019. p. 261-262.

[36] Prova disso é o famoso caso envolvendo a empresa Cambridge Anlalytica, ocasião em que os dados utilizados foram publicados voluntariamente pelos titulares e, mesmo assim, foram utilizados de forma a ferir o seu direito ao livre desenvolvimento da personalidade. Vide: CADWALLADR, Carole; GRAHAM-HARRISON, Emma. **Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach**. 17 mar. 2018. Disponível em: <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>. Acesso em 25 set. 2021.

[37] BRASIL. Supremo Tribunal Federal. ADIs nº 6.387, 6.388, 6.389, 6.390 e 6.393. Relatora: Ministra Rosa Weber. 7 de maio de 2020. Disponível em: <http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895165> Acesso em 21 set. 2021.

considerações íntimas, como crenças, posições políticas e opiniões[38]. A legislação aplica-se a toda a variedade de informações pessoais indistintamente[39], independendo da maneira pela qual foram recolhidas ou do suporte em que se encontram armazenadas (se físico ou digital).

2.1.2 Relacionada

É imprescindível a identificação de quais ligações são relevantes e como distingui-las[40]. Regra geral, para que a informação seja considerada dado pessoal, ela deve ser relativa a uma pessoa, isso é, deve ser *sobre* a pessoa[41]. Logo, informações relacionadas a situações não subjetiváveis, como, por exemplo, objetos, eventos e animais, não podem ser consideradas dados pessoais, estando, assim, excluídas do campo de aplicação da LGPD.

Contudo, é possível que, quando combinada com uma informação relativa a uma pessoa, a informação relacionada à realidade não subjetivável seja tratada como dado pessoal[42] – por exemplo: (i) um computador com número IP A conectou-se à rede X; (ii) o computador pertence a Carlos. A informação (i) é relacionada a um objeto, não sendo, pois, dado pessoal. Todavia, quando conjugada à informação (ii), a informação (i) converte-se em dado pessoal, já que passa a revelar aspecto referente à determinada pessoa – Carlos conectou-se à rede X por meio do seu computador que detém número IP A.

Da mesma forma, ainda que não combinada com uma informação pessoal, dados factuais poderão ser reconduzidos ao universo dos dados pessoais a depender do contexto em que forem analisados. Traz-se, a título elucidativo, o exemplo referido pelo GT 29 no Parecer 4/2007, relativo ao valor de uma casa[43]. Tal informação refere-se, indiscutivelmente, a um objeto. As regras de proteção de dados não serão aplicadas quando esta informação for

---

[38] A informação pode ter natureza objetiva – ex.: A trabalha na empresa X; ou subjetiva – ex.: A é um trabalhador dedicado.

[39] Todas as modalidades de dados recebem igual proteção da Lei, todavia, há aqueles que, por sua natureza, devem receber maior grau de atenção dos agentes de tratamento quando do seu processamento. São eles os dados sensíveis (revelam aspectos relativos à origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, saúde, vida sexual, etc.) e os dados relativos a crianças e adolescentes. Vide: BRASIL. Governo Federal. O que são dados sensíveis, de acordo com a LGPD. Disponível em: <https://www.serpro.gov.br/lgpd/menu/protecao-de-dados/dados-sensiveis-lgpd>. Acesso em 21 set. 2021.

[40] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 09. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

[41] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 110.

[42] *Ibidem*.

[43] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 10. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

utilizada apenas para ilustrar o nível dos preços imobiliários em uma determinada zona. No entanto, em certas circunstâncias, esta informação poderá ser também considerada como dado pessoal: a casa é um ativo do seu proprietário, podendo servir para, por exemplo, determinar o alcance da capacidade aquisitiva desta pessoa para pagar determinados impostos. Neste contexto, esta informação deverá ser considerada como dado pessoal.

Consoante proposto pelo GT 29, a expressão "relacionada a" pode abranger três situações distintas, as quais configuram elementos alternativos e não cumulativos: (i) conteúdo; (ii) finalidade; e (iii) resultado. A primeira delas constitui a regra geral antes mencionada, na qual a informação é sobre determinada pessoa; o indivíduo é o objeto da informação. Este elemento independe do contexto e do objetivo do responsável pelo tratamento dos dados para que a informação seja considerada pessoal. Estão incluídos neste grupo, por exemplo, as fichas médicas, nas quais todas as informações ali contidas dizem respeito a determinado sujeito, qual seja o paciente. Na segunda situação, a pessoa não é o objeto da informação. O que acontece é que os dados são utilizados para uma finalidade específica que permite avaliar, tratar de determinada forma ou influenciar o comportamento de um indivíduo. É o caso de dados que constem em um registro de tempo de trabalho alusivos ao período de serviço e ao período de descanso de determinado trabalhador[44]. Por fim, as situações que contenham elemento de resultado referem-se àquelas que não têm o indivíduo como objeto (conteúdo) tampouco visam avaliá-lo ou influenciá-lo (finalidade), mas que, mesmo assim, provoquem impacto nos direitos ou interesses de determinada pessoa. Para melhor ilustração, cita-se o exemplo trazido pelo GT 29: uma empresa de táxis instala um sistema de localização por satélite para determinar, em tempo real, a posição dos carros disponíveis. Apesar de o objeto das informações colhidas ser os carros e a finalidade da coleta ser o fornecimento de melhor serviço aos usuários, tal tecnologia permite controlar o desempenho dos motoristas e verificar se respeitam os limites de velocidade. As informações colhidas tem, portanto, impacto considerável nestes trabalhadores, de modo que podem ser consideradas dados pessoais merecedores de proteção jurídica[45].

Ressalta-se que a expressão "relacionada a" pode se referir a uma pessoa ou a um grupo de pessoas. O titular não precisa ser representado de maneira individualizada pela informação para que ela seja considerada pessoal e o seu tratamento seja sujeito às regras de

---

[44] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 112.

[45] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 11-12. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

proteção de dados. Exemplos desta situação são as informações relativas a um grupo étnico ou racial – tais dados são compartilhados por um grupo de titulares e, além disso, os indivíduos podem ser visados como um membro desse grupo específico[46].

2.1.3 Pessoa Natural

Para que o dado seja considerado pessoal, ele deverá deter um vínculo objetivo com determinada pessoa natural, revelando algo sobre ela[47]. O objeto do regime de proteção estabelecido pela LGPD não é o dado em si, mas, justamente, a pessoa natural à qual ele se relaciona[48], independentemente da sua nacionalidade ou local de sua residência. Pode-se dizer que, neste aspecto, o direito à proteção de dados pessoais é universal, não se restringindo a nacionais ou residentes de um determinado país[49], sendo, pois, aplicável a todas as pessoas naturais, ou seja, a todos os seres humanos.

Pessoas jurídicas não são titulares de dados pessoais[50], tampouco coisas ou animais (realidades jurídicas não subjetiváveis). A informação relativa à pessoa jurídica apenas poderá ser merecedora de tutela da LGPD quando se relaciona, direta ou indiretamente, com uma pessoa natural – por exemplo, quando o nome do sócio consta na razão social de uma sociedade empresarial, as informações referentes à saúde financeira desta poderão ser associadas à figura daquele[51-52]. Há, contudo, discussões na doutrina nacional e estrangeira acerca da possibilidade de o direito à proteção de dados pessoais ser estendido às pessoas jurídicas, baseando-se no entendimento de que os atos negociais praticados por estes entes

---

[46] PAGALLO, Ugo. The Group, the Private, and the Individual: A New Level of Data Protection? In: FLORIDI, Luciano; TAYLOR, Linnet; VAN DER SLOOT, Bart (coord.). **Group Privacy**: New Challenges of Data Technologies. Dordrecht: Springer International Publishing, 2017. *E-book*. p. 160.

[47] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais:** elementos da formação da Lei Geral de Proteção de Dados. 2 ed. São Paulo: Thomson Reuters Brasil, 2019. p. 139.

[48] MAYER-SCHÖNBERGER Viktor. Generational Development of Data Protection in Europe. In: AGRE, Philip E.; ROTENBERG, Marc (orgs.). **Technology and Privacy:** The New Landscape. Cambridge: The MIT Press, 1997. p. 219.

[49] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 23. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

[50] Recital 14: The protection afforded by this Regulation should apply to natural persons, whatever their nationality or place of residence, in relation to the processing of their personal data. This Regulation does not cover the processing of personal data which concerns legal persons and in particular undertakings established as legal persons, including the name and the form of the legal person and the contact details of the legal person. EUROPEAN UNION. Recital 14. **General Data Protection Regulation – GDPR**. Disponível em: <https://www.privacy-regulation.eu/en/recital-14-GDPR.htm>. Acesso em 25 set. 2021.

[51] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 113.

[52] BUCHNER, Benedikt; KÜHLING, Jürgen (org.). **Datenschutz-Grundverordnung BDSG** – Kommentar. 3. ed. München: C.H.Beck, 2020, Art. 4 Abs. 1 Rn. 3-7. *E-book*.

serão atribuídos à própria empresa, e não ao sócio, de modo que os dados oriundos de contratações (eletrônicas ou não) são prolongamentos da identidade econômica e social daquela, merecendo proteção ainda que relativos a pessoas jurídicas[53-54]. Defende-se que tal proteção teria como fundamento a salvaguarda dos direitos econômicos destes entes, e não de direitos fundamentais, como o é em relação às pessoas naturais[55].

No tocante à pessoa natural, cabe delimitar se a tutela é estendida ou não aos nascituros e às pessoas falecidas. Ao contrário do que acontece no Regulamento Europeu de Proteção de Dados (*General Data Protection Regulation – GDPR*), em que se exclui, expressamente, do escopo de proteção os dados referentes a pessoas falecidas (Considerandos 27, 158 e 160), não existe, na LGPD, qualquer disposição neste sentido. Contudo, em se considerando a previsão do art. 6º do Código Civil que dispõe que "a existência da pessoa natural termina com a morte", possível a presunção de que o tratamento de dados relativos a pessoas falecidas não se submete às regras da LGPD, já que aquelas não mais são consideradas pessoas naturais. Há, todavia, uma corrente contrária que defende a não exclusão de dados relativos a pessoas falecidas do escopo de aplicação da Lei, em razão do disposto no art. 12 do Código Civil, em que se assegura a adoção de medidas para a proteção dos direitos da personalidade da pessoa falecida, pelo cônjuge sobrevivente, ou qualquer parente em linha reta, ou colateral até o quarto grau. Cumpre ressaltar que, assim como ocorre com os dados relativos a pessoas jurídicas, aqueles que forem relacionados a pessoas falecidas, mas que acabam revelando algum aspecto de uma pessoa natural, merecem, indubitavelmente, a proteção oferecida pela LGPD. Por exemplo, a informação de que o falecido tinha sofrido de uma doença hereditária não é um dado protegido em relação àquele. No entanto, em relação a seu descendente vivo, o dado pode ser categorizado como "pessoal" se a informação sugerir que este último também é afetado pela doença hereditária[56].

---

[53] MAGALHÃES, Rodrigo Almeida; DIVINO, Sthéfano Bruno Santos. A proteção de dados da pessoa jurídica e a Lei 13.709/2018: reflexões à luz dos direitos da personalidade. **Scientia Iuris**, v. 23, n. 2, p. 74-90, jul. 2019. p. 43.

[54] Rodrigo Magalhães e Sthéfano Divino exemplificam esta concepção: "Com apenas um *click* na caixa de diálogo (Li e aceito os termos de serviços e as políticas de privacidade), terá o controlador e/ou operador o relevo necessário ao início das atividades de processamento de dados. O reconhecimento desta situação jurídica independe da categorização de pessoa. Se jurídica ou se natural, ocorrerá de qualquer forma.". Vide: MAGALHÃES, Rodrigo Almeida; DIVINO, Sthéfano Bruno Santos. A proteção de dados da pessoa jurídica e a Lei 13.709/2018: reflexões à luz dos direitos da personalidade. **Scientia Iuris**, v. 23, n. 2, p. 74-90, jul. 2019. p. 43.

[55] CHAMOUX, Jean Pierre. Data Protection in Europe - The Problem of the Physical Person and the Legal Person. **Journal of Media Law and Practice,** v. 2, n. 1, p. 70-83, 1981. p. 74.

[56] BUCHNER, Benedikt; KÜHLING, Jürgen (org.). **Datenschutz-Grundverordnung BDSG** – Kommentar. 3. ed. München: C.H.Beck, 2020, Art. 4 Abs. 1 Rn. 3-7. *E-book*.

Em relação aos dados relativos a nascituros – tais como ecografias e análises do líquido amniótico – a LGPD nem o GDPR[57] referem se esses são ou não abrangidos pelo regime de proteção de dados pessoais. O Código Civil Brasileiro prevê a proteção dos direitos do nascituro desde a sua concepção[58], assim, ainda que este não tenha personalidade civil, os seus direitos encontram-se resguardados pela lei. Em sendo o direito à proteção de dados pessoais um destes direitos, poder-se-ia admitir que nascituros fossem titulares de dados pessoais, merecendo, pois, proteção[59]. Discute-se, também, de quem é a titularidade de tais informações: da genitora ou do próprio nascituro[60].

2.1.4 Identificada ou Identificável

O vínculo estabelecido entre o dado e o seu titular pode se dar de forma imediata/precisa – neste caso, a pessoa natural é identificada; ou de forma mediata/imprecisa – a pessoa natural é identificável[61-62]. A LGPD não especifica, de forma expressa, o que caracteriza cada um destes elementos. Deste modo, diante da omissão da Lei brasileira, pode-se recorrer a um diálogo normativo com a regulação que em muito influenciou a LGPD: o

[57] A decisão referente à aplicação das regras de proteção de dados aos nascituros foi delegada, pelo legislador europeu, aos Estados-Membros. Cada Estado o define a depender do posicionamento geral do seu sistema jurídico sobre a proteção dos nascituros. Vide: EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 24. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

[58] Art. 2º. A personalidade civil da pessoa começa do nascimento com vida; mas a lei põe a salvo, desde a concepção, os direitos do nascituro. BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002**. Código Civil. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm>. Acesso em 01 set. 2021.

[59] Poder-se-ia determinar esta questão a partir da análise das teorias relacionadas à titularidade, pelo nascituro, de direitos de personalidade. Todavia, por não ser o foco da presente monografia, o assunto não será aprofundado.

[60] Sobre a temática, vide: PORMEISTER, Kart; DROZDZOWSKI, Lukasz. Protecting the Genetic Data of Unborn Children: A Critical Analysis. **European Data Protection Law Review (EDPL)**, v. 4, n. 1, p. 53-64, 2018.

[61] BIONI, Bruno. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, jul. 2015. p. 17.

[62] Bruno Bioni propõe uma diferenciação entre o conceito de dado pessoal que apenas admite a qualificação da pessoa como identificada, o qual ele denomina de "conceito reducionista", daquele que admite que ela seja meramente identificável, denominado por ele de "conceito expansionista". Este último, de acordo com o autor, é o adotado pela LGPD, já que, além de admitir que a pessoa seja identificada, a Lei também reconhece a possibilidade de esta ser identificável. Vide: BIONI, Bruno. **Proteção de dados pessoais**: a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019. p. 66: "De forma sistemática, o vocabulário para prescrever tal definição é composto por palavras que restringem ou largam o *gargalo* dessa proteção. Há uma bipartição do seu léxico que ora retrai (reducionista), ora expande (expansionista), a moldura normativa de uma lei de proteção de dados pessoais.". No mesmo sentido: SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII problem: privacy and a new concept of personally identifiable information. **New York University Law Review**, v. 86, p. 1814-1894, 2011. p. 1871-1877.

Regulamento Europeu de Proteção de Dados. O legislador europeu define o que se entende por "identificável" no próprio artigo em que se conceitua "dado pessoal":

> Art. 4 - (1) 'Dado pessoal' é toda informação relativa a uma pessoa singular identificada ou identificável ('titular dos dados'); é considerada identificável uma pessoa natural que possa ser identificada, direta ou indiretamente, em especial por referência a um identificador, como, por exemplo, nome, número de identificação, dados de localização, identificadores por via eletrônica ou a um ou mais elementos específicos da identidade física, fisiológica, genética, mental, econômica, cultural ou social dessa pessoa natural; [...].[63]

O Considerando 26 ainda complementa:

> [...] Para determinar se uma pessoa natural é identificável, deve-se levar em conta todos os meios suscetíveis de serem razoavelmente utilizados, tais como a distinção, quer pelo responsável pelo tratamento, quer por outra pessoa, para identificar direta ou indiretamente a pessoa natural.[64]

Para compreender a definição de "identificável", se faz necessário, em primeiro momento, ter em mente o que significa o termo "identificada". A pessoa à qual o dado se relaciona é considerada "identificada" quando esse possibilita o conhecimento e a individualização daquela imediatamente. Neste caso, o processo de identificação se dá por meio de identificadores, isso é, informações especiais que "têm uma relação especialmente privilegiada e próxima com a pessoa em causa"[65]. Os identificadores são classificados como diretos quando constituem o primeiro sinal distintivo da individualidade de um sujeito[66], não

---

[63] EUROPEAN UNION. Regulation (EU) 2016/679. **General Data Protection Regulation – GDPR**. Disponível em: <https://gdpr-info.eu/>. Acesso em 25 set. 2021.
[64] EUROPEAN UNION. Recital 26. Regulation (EU) 2016/679. **General Data Protection Regulation – GDPR**. Disponível em: <https://gdpr-text.com/pt/read/recital-26/>. Acesso em 21 set. 2021.
[65] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 13. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.
[66] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.2.

sendo necessários dados adicionais para viabilizar a identificação[67] – por exemplo, o nome completo de uma pessoa (prenome + sobrenome), desde que seja único[68], permite diferenciá-la de forma inequívoca de todos os demais indivíduos. Por sua vez, são chamados de identificadores indiretos aqueles que, em sendo combinados entre si, formam uma combinação única capaz de identificar determinado sujeito[69]. São exemplos de identificadores indiretos o local de nascimento, a raça, a religião, o peso, etc[70].

Quando a identificação não acontece de forma imediata, mas, sim, mediata, a pessoa é considerada identificável. Em outras palavras, nos casos em que, *prima facie*, os identificadores disponíveis não permitem isolar determinada pessoa, mas, se combinados a outras informações, independentemente de essas estarem ou não à disposição do responsável pelo tratamento, possibilitarem que ela seja distinguida, então o titular é identificável[71]. Há um potencial de o sujeito se tornar identificado, o que pode se dar direta ou indiretamente. O endereço de IP dinâmico, por exemplo, é uma informação relativa a uma pessoa identificável, na medida em que não revela imediatamente a identidade da pessoa natural proprietária do computador a partir do qual se efetua a consulta de um *site* na Internet[72], mas ela pode ser identificada "mediante requerimento judicial de acesso a registros de conexão e dados cadastrais armazenados pelos respectivos provedores de conexão"[73]. Conforme disposto no Parecer 4/2007 sobre o conceito de dados pessoais elaborado pelo WP29, a pessoa é identificável quando é possível a sua identificação por meio de uma "combinação de critérios significativos que permitem o reconhecimento da pessoa por eliminação dos elementos do

---

[67] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 121.

[68] Pode acontecer que em uma base de dados haja mais de uma pessoa com o mesmo nome, neste caso, são necessários dados adicionais, de modo que a pessoa é, portanto, identificável.

[69] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 14. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

[70] EUROPEAN UNION. Collaboration in Research and Methodology for Official Statistics - CROS. **Indirect identification**. Disponível em: <https://ec.europa.eu/eurostat/cros/content/indirect-identification_en>. Acesso em 16 out. 2021.

[71] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 14. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

[72] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 121.

[73] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.2.

grupo a que ela pertence (idade, profissão, local de residência, etc.)"[74]. Isso se torna mais claro quando da análise de um banco de dados relacional estruturado em tabela:

Quadro 1 – Banco de dados relacional.

NOME	CPF	CEP	IDADE	SEGMENTAÇÃO
Carlos Silva	219.755.046-08	91978-055	16	Jovem jogador de tênis
Carlos Silva	354.477.042-06	42900-020	17	Jovem sedentário
Carlos Friedrich	219.756.877-84	91978-070	21	Jovem fisiculturista
Carlos Freitas	354.476.590-33	42368-678	30	Adulto ciclista

Fonte: Adaptado a partir de BIONI, Bruno. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, jul. 2015. p. 20.

No exemplo supra, o nome completo dos dois primeiros indivíduos não é suficiente para identificá-los (e diferenciá-los) quando da análise do banco de dados, porquanto se tratam de homônimos. Contudo, por meio da combinação do nome com um identificador único[75], como, por exemplo, o CPF, cria-se uma combinação única que possibilita a individualização dos "Carlos Silvas" de forma exata e inequívoca, tornando-os identificados[76]. Porém, se não houvesse as colunas referentes ao CPF e ao CEP, tornar-se-ia mais difícil a individualização de cada um deles, de modo que seriam considerados identificáveis. A identificação foi dificultada, mas não descartada, já que haveria um potencial de os titulares se tornarem identificados, caso fosse possível a coleta de mais informações por meio da utilização de meios razoáveis[77].

Para determinar o grau de identificabilidade de um dado, demanda-se a realização de uma análise contextual de onde ele está inserido[78], verificando a necessidade de combiná-lo a

[74] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 13. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.
[75] Identificador único é uma cadeia numérica ou alfanumérica que está associada a uma única pessoa dentro de um determinado sistema, permitindo distingui-lo inequivocamente dos demais indivíduos. Vide: WIGMORE, Ivy. **Unique Identifier (UID)**. Disponível em: <https://internetofthingsagenda.techtarget.com/definition/unique-identifier-UID>. Acesso em 11 out. 2021.
[76] BIONI, Bruno. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, jul. 2015. p. 20-21.
[77] BIONI, Bruno. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, jul. 2015. p. 20-21.
[78] BIONI, Bruno. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, jul. 2015. p. 18.

outras informações existentes para individualizar o seu titular. Isso porque na medida em que determinados identificadores são suficientes para obter a individualização da pessoa, outros necessitam de informações adicionais, a depender do contexto da situação em causa. Inclusive, o mesmo dado pode se relacionar a uma pessoa identificada e, em outra situação, a uma pessoa identificável. Um sobrenome recorrente no Brasil (como é o caso de "Silva", utilizado no exemplo acima) não será suficiente para identificar uma pessoa – isto é, para distingui-la – de toda a população do país, por outro lado, é provável que permita a identificação de um aluno em determinada sala de aula[79]. Outra situação possível é um identificador que não é considerado único, sendo, pois, informação acessória, tal como "a mulher com um casaco vermelho" ser o bastante para identificar alguém de entre um grupo de clientes aguardando no rol de entrada de um escritório[80]. Assim, determinar se a pessoa a que a informação é relativa está identificada ou é identificável depende das circunstâncias do contexto em análise e da autossuficiência da informação detida[81].

2.2 A DUALIDADE ENTRE OS CONCEITOS DE DADO PESSOAL E DE DADO ANONIMIZADO

Conforme já abordado, o conceito de dado pessoal é imprescindível na interpretação do alcance normativo[82] da LGPD, já que esta tem como objetivo a tutela das pessoas naturais, identificadas ou identificáveis, que possuem vínculo com dados visados para tratamento. Quando a informação não possuir tal vínculo, estará fora do âmbito da tutela jurídica oferecida pela Lei. Pode-se dizer, pois, que tal "conceituação também estabelece, consequentemente, a linha divisória do que é e não é informação pessoal"[83]. Apenas os dados que atraem o qualificador pessoal possuem relevância jurídica no âmbito de aplicabilidade das

---

[79] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 13. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.
[80] *Ibidem*.
[81] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 122.
[82] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.2.
[83] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.2.

regras de proteção de dados[84]. Por conta disso, é recorrente a afirmação de que "a antítese do conceito de dado pessoal é o conceito de dado anônimo"[85]. O dado anônimo ou anonimizado[86] é aquele que é incapaz de revelar a identidade de uma pessoa[87], pois não possui vínculo com nenhum sujeito ou possui vínculo com sujeito indeterminado[88]. Sobre esta categoria não incide a proteção concedida pela LGPD, justamente porque o tratamento destes dados não afeta a esfera pessoal de pessoa natural, tampouco representa prolongamento (mediato ou imediato) desta.

A LGPD define dado anonimizado como "dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento"[89]. A identificação do titular é, portanto, o elemento central para diferenciar um dado pessoal de um dado anonimizado. O dado não pode, em tese, apresentar qualquer grau de identificabilidade para que seja considerado anônimo. Por exemplo, a idade de determinado indivíduo não é um identificador único, pois não permite a identificação imediata e inequívoca do sujeito, tampouco um dado anônimo, pois, quando combinada com outros dados, tais como endereço e gênero, possibilitaria chegar ao titular. Assim sendo, se o titular do dado for (ainda que em grau reduzido) identificável, ou se puder ser conectado direta ou indiretamente ao dado, tanto o controlador[90] quanto o operador[91] deverão observar os princípios da proteção de dados pessoais quando do tratamento desta informação.

---

[84] BIONI, Bruno. **Proteção de dados pessoais**: a função e os limites do consentimento. 1ª ed. Rio de Janeiro: Forense, 2019. p. 68.
[85] *Ibidem*. p. 70.
[86] Alguns autores diferenciam o dado anônimo do dado anonimizado. Enquanto aquele seria o dado que nunca teve qualquer tipo de vínculo com um sujeito, este seria o dado pessoal que passou por processo de anonimização, tornando indeterminado o indivíduo ao qual o dado se vincula. Considerando a nomenclatura adotada na LGPD e na maior parte da doutrina nacional, no presente trabalho, não será adotada a referida diferenciação, de modo que os termos serão usados como sinônimos.
[87] BIONI, Bruno. **Proteção de dados pessoais**: a função e os limites do consentimento. 1ª ed. Rio de Janeiro: Forense, 2019. p. 70.
[88] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2 ed. São Paulo: Thomson Reuters Brasil, 2019. p. 140.
[89] Art. 5º. Para os fins desta Lei, considera-se: [...] III - dado anonimizado: dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.
[90] Art. 5º. Para os fins desta Lei, considera-se: [...] VI - controlador: pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.
[91] Art. 5º. Para os fins desta Lei, considera-se: [...] VII - operador: pessoa natural ou jurídica, de direito público ou privado, que realiza o tratamento de dados pessoais em nome do controlador; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

Apesar da dualidade existente, a linha que separa o dado pessoal do dado anônimo é bastante tênue na prática. Um dado pessoal pode ser transmudado em um dado anonimizado, assim como um dado anonimizado pode passar a ser considerado dado pessoal. Por meio do processo de anonimização, o dado pessoal pode perder o vínculo antes estabelecido com o seu titular, tornando-se anônimo. Contudo, o avanço da tecnologia e o crescente fornecimento de dados facilita a inferência de informações pessoais a partir da ligação de base de dados ostensivamente anonimizados[92]. Dada a possibilidade destas mutações, é importante, além de diferenciar dado pessoal de dado anonimizado, o fazer em relação a outros dois tipos comumentemente confundidos a este: os dados pseudonimizados e os dados criptografados.

2.2.1 Dado anonimizado vs. Dado pseudonimizado

A LGPD não traz o conceito de dado pseudonimizado em seu artigo 5º, dispositivo em que constam as principais definições necessárias para a compreensão do conteúdo da Lei. Todavia, no artigo 13, o processo de pseudonimização é citado no *caput* e, posteriormente, definido no §4º:

> §4º Para os efeitos deste artigo, a pseudonimização 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.[93]

A leitura atenta do supracitado dispositivo permite constatar que a pseudonimização é uma modalidade de tratamento de dados pessoais. Logo, o dado pseudonimizado não configura uma terceira categoria de dados, tampouco pode ser confundido com um dado anonimizado[94]. Com a adoção da pseudonimização, o responsável pelo tratamento detém, em um banco de dados em separado, informações adicionais que permitem identificar o titular.

---

[92] FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020. p. 34.

[93] BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[94] Nesse sentido: FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020. p. 21: "There is an ongoing debate regarding the implications of Article 4(5), in particular, whether the provision gives rise to a third category of data beyond those of personal and anonymous data. A literal interpretation reveals, however, that Article 4(5) GDPR deals with a method, not an outcome of data processing. Pseudonymization is the 'processing' of personal data in such a way that data can only be attributed to a data subject with the help of additional information. This underlines that pseudonymized data remains personal data […]".

Os dados pessoais são apenas "camuflados", ou seja, "substitui-se um atributo (tipicamente um atributo único) em um registro por outro"[95]. Trata-se, pois, de uma medida de segurança da informação, em que se diminui a possibilidade de identificação, sem descartá-la - ainda é possível identificar, indiretamente, a pessoa natural titular do dado. Novamente, recorre-se à base de dados relacional estruturada em tabela como recurso didático:

Quadro 2 – Base de dados relacional pseudonimizada.

NOME	CPF	IDADE	SEGMENTAÇÃO	NÚMERO DE REFERÊNCIA
[redacted]	[redacted]	16	Jovem jogador de tênis	A6153BHL9
[redacted]	[redacted]	17	Jovem sedentário	HA7166236
[redacted]	[redacted]	21	Jovem fisiculturista	UH9147NT9
[redacted]	[redacted]	30	Adulto ciclista	WY791O240

Fonte: Adaptado a partir de EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 201. p. 20. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

A técnica de pseudonimização aplicada na tabela acima chama-se *hashing*. Nela, o controlador gera uma codificação aleatória por meio dos identificadores excluídos da base de dados (Nome e CPF) que funciona como um "pseudônimo" ao titular das informações coletadas. Em separado, mantém-se uma listagem das identidades com seus respectivos pseudônimos, o que torna as pessoas que titulam os dados pseudonimizados identificáveis.

Aos dados pseudonimizados aplicam-se as normas de proteção de dados pessoais, uma vez que a possibilidade de identificação do titular é mantida pelo controlador, razão pela qual os dados mantém o qualificador 'pessoal'[96]. Por conta disso, muitos questionam qual seria a vantagem da adoção do procedimento de pseudonimização, já que este não afasta os deveres e obrigações advindos da LGPD. A resposta é a mitigação dos riscos provenientes do processamento de dados. No Regulamento Europeu (GDPR), algumas normas aplicadas aos

[95] "Pseudonymisation consists of replacing one attribute (typically a unique attribute) in a record by another." EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 20 (tradução livre). Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

[96] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 18. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021: "Retraceably pseudonymised data may be considered as information on individuals which are indirectly identifiable..".

dados pseudonimizados são mais flexíveis em relação àquelas aplicáveis aos dados pessoais[97], o que é justificado pelo legislador europeu pelo fato de os riscos serem consideravelmente reduzidos. Esta flexibilização funciona, também, como um incentivo à adoção da pseudonimização ao tratar dados pessoais. Na LGPD, não há previsão expressa de que as regras aplicáveis a dados pseudonimizados seriam mais brandas, contudo, no *caput* do artigo 13, a Lei determina que, para fins de pesquisa com dados pessoais sensíveis, a pseudonimização deve ser aplicada sempre que possível[98].

Em determinadas atividades, sejam elas realizadas por entes públicos ou privados, a identificação do titular dos dados coletados precisa ser mantida para que a finalidade daquelas seja alcançada com plena efetividade. Nestas, a adoção da anonimização dos dados tornar-se-ia um obstáculo, sendo a solução o emprego da pseudonimização. Neste ponto, traz-se como exemplo a plataforma Fala.BR que consiste "em um canal integrado para encaminhamento de manifestações (acesso a informação, denúncias, reclamações, solicitações, sugestões, elogios e simplifique) a órgãos e entidades do poder público"[99]. Nas modalidades "denúncia" e "reclamação", o manifestante poderá realizá-las de forma anônima, sem necessidade de cadastro no sistema. Todavia, neste caso, a ouvidoria não poderá responder à manifestação, tampouco solicitar complementação ou prestar orientações[100]. Da mesma forma, o denunciante perde a possibilidade de fazer o acompanhamento da sua denúncia (se foi enviada para apreciação, se foi arquivada…), o que pode ser de seu interesse. Portanto, se os dados forem anônimos, extingue-se a interação entre os órgãos e o cidadão, e, consequentemente, o serviço prestado torna-se menos efetivo. Assim, para não limitar a prestação do serviço ao cidadão e preservar a sua confiança no Estado (o que é especialmente

---

[97] Article 6. Lawfulness of processing. [...] 4. Where the processing for a purpose other than that for which the personal data have been collected is not based on the data subject's consent or on a Union or Member State law which constitutes a necessary and proportionate measure in a democratic society to safeguard the objectives referred to in Article 23(1), the controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, take into account, inter alia: [...] (e) the existence of appropriate safeguards, which may include encryption or pseudonymisation. EUROPEAN UNION. Recital 14. **General Data Protection Regulation – GDPR**. Disponível em: <https://www.privacy-regulation.eu/en/recital-14-GDPR.htm>. Acesso em 25 set. 2021.

[98] Art. 13. Na realização de estudos em saúde pública, os órgãos de pesquisa poderão ter acesso a bases de dados pessoais, que serão tratados exclusivamente dentro do órgão e estritamente para a finalidade de realização de estudos e pesquisas e mantidos em ambiente controlado e seguro, conforme práticas de segurança previstas em regulamento específico e que incluam, sempre que possível, a anonimização ou pseudonimização dos dados, bem como considerem os devidos padrões éticos relacionados a estudos e pesquisas. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[99] BRASIL. Governo Federal. Perguntas e Respostas. Disponível em: <https://www.gov.br/ouvidorias/pt-br/central-de-conteudos/perguntas-frequentes-2019#resp1>. Acesso em 16 out. 2021.

[100] Webinário: Anonimização e pseudonimização de dados pessoais - conversas sobre a LGPD e o papel da Ouvidoria. 1 vídeo (162 min). Publicado pelo canal Controladoria-Geral da União – CGU. Disponível em: <https://www.youtube.com/watch?v=8AVj0wmzFRs>. Acesso em: 29 jun. 2021.

importante na ouvidoria pública), a técnica adotada em tal plataforma tem sido, em regra, a pseudonimização[101]. O denunciante que quiser preservar o canal de interação poderá optar por realizar o cadastro no sistema, sem ter seus dados pessoais acessados indevidamente, pois contará, de imediato, com a aplicação de medidas de proteção, quais sejam, a pseudonimização das informações fornecidas[102]. Marcos Lindenmayer, Chefe de Gabinete da Ouvidoria-Geral da União, concede mais detalhes sobre a adoção do procedimento no âmbito da plataforma Fala.BR:

> Quando o denunciante realiza uma denúncia anônima, não é inserido nenhum dado dentro do sistema que permitiria uma interlocução do Poder Executivo Federal com ele. Por uma questão de proteção, o denunciante não tem acesso aos desdobramentos da denúncia. A partir de 2019, com o Decreto 10.153, aquele que se identificou tem sua identidade protegida "por *default*", ou seja, a pessoa que realiza a denúncia passa a contar com as medidas de proteção à sua identidade de imediato, sem necessidade de ter que tomar alguma atitude específica para solicitar a proteção.[103]

A pseudonimização vem, portanto, como uma alternativa à anonimização para que as empresas e órgãos públicos intensifiquem a segurança das informações a eles fornecidas, sem que isso afete a efetividade de suas atividades. Apesar de a diferenciação ser feita em sede legal, tanto na LGPD quanto na GDPR[104], há na doutrina quem entenda que a pseudonimização seria uma técnica de anonimização, todavia, esse entendimento é superável porquanto para que os dados sejam considerados pseudonimizados, eles devem ser, justamente, reidentificáveis[105].

---

[101] Art. 6º. [...] § 4º A unidade de ouvidoria responsável pelo tratamento da denúncia providenciará a sua pseudonimização para o posterior envio aos órgãos de apuração competentes, observado o disposto no § 2º. BRASIL. **Decreto nº 10.153, de 3 de dezembro de 2019**. Disponível em: <https://www.in.gov.br/web/dou/-/decreto-n-10.153-de-3-de-dezembro-de-2019-231274119>. Acesso em 17 out. 2021.

[102] Webinário: Anonimização e pseudonimização de dados pessoais - conversas sobre a LGPD e o papel da Ouvidoria. 1 vídeo (162 min). Publicado pelo canal Controladoria-Geral da União – CGU. Disponível em: <https://www.youtube.com/watch?v=8AVj0wmzFRs>. Acesso em: 29 jun. 2021.

[103] *Ibidem*.

[104] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book*. p. RB-8.2.

[105] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 19. Disponível em: <https://ec.europa.eu/justice/article-

2.2.2 Dado anonimizado vs. Dado criptografado

A criptografia, uma das técnicas de segurança da informação mais antigas que se tem conhecimento[106], surgiu com a finalidade de manter a confidencialidade de mensagens, especialmente, no âmbito militar[107]. Etimologicamente, a palavra 'criptografia' tem como significado "a ciência da escrita secreta que tem como objetivo esconder o significado de mensagens"[108]. Apesar de a sua criação ter sido motivada pelo sigilo, hoje, a criptografia não mais se restringe a esta finalidade, constituindo uma forma de proteção de direitos – liberdade de expressão, privacidade, proteção de dados pessoais – e de garantir a integridade do conteúdo de dados e mensagens (textos, vídeos, imagens, áudios) tanto em face de um "intruso" quanto do próprio provedor do serviço[109]. Isso foi amplamente reconhecido pelo Supremo Tribunal Federal quando do julgamento da ADPF 403 e da ADI 5527, em maio de 2020[110], ocasião em que foi analisada a aplicação da criptografia ponta a ponta no âmbito do aplicativo *WhatsApp*. Em seu voto, o Ministro Luiz Edson Fachin, relator da ADPF 403/SE, ressaltou:

> [...] a criptografia protege os direitos dos usuários da internet, garantindo a privacidade de suas comunicações, e que, portanto, é do interesse do Estado brasileiro encorajar as empresas e as pessoas a

---

29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021. No mesmo sentido: EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 03. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021: "[...] pseudonymisation is not a method of anonymisation. It merely reduces the linkability of a dataset with the original identity of a data subject, and is accordingly a useful security measure.".

[106] Vide: PAAR, Christof; PELZL, Jan. **Understanding cryptography**: a textbook for students and practitioners. Londres: Springer, 2010. p. 2: "Cryptography seems closely linked to modern electronic communication. How- ever, cryptography is a rather old business, with early examples dating back to about 2000 B.C., when non-standard "secret" hieroglyphics were used in ancient Egypt*"*.

[107] ARANHA, Diego F. O que é criptografia fim a fim e o que devemos fazer a respeito? In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book*. p. RB-1.1.

[108] "Cryptography is the science of secret writing with the goal of hiding the meaning of a message." PAAR, Christof; PELZL, Jan. **Understanding cryptography**: a textbook for students and practitioners. Londres: Springer, 2010. p. 3 (tradução livre).

[109] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book*. p. RB-8.3.

[110] O julgamento conjunto da ADPF 403 e da ADI 5527 ainda está em andamento, tendo sido proferido somente dois votos, quais sejam, o do Min. Edson Fachin (relator da ADPF 403) e o da Min.ª Rosa Weber (relatora da ADI 5527).

> utilizarem a criptografia e manter o ambiente digital com a maior segurança possível para os usuários.[111]

A adoção de técnicas criptográficas como medida contra riscos de incidentes de segurança em bases de dados é muito comum em diversos serviços prestados para a sociedade - *e-commerce*, serviços bancários, correio eletrônico, assinatura digital, etc. Tal procedimento funciona por meio do uso de uma chave criptográfica, isso é, de um conjunto de algoritmos, simétricos ou assimétricos[112], projetado para ser absolutamente único[113]. As informações que passam pela chave são codificadas, dando origem a dados criptografados (diga-se, ilegíveis). A reversão somente pode ser realizada mediante o conhecimento da referida chave, que estabelece a correspondência entre o código e as informações codificadas[114]. Logo, somente os possuidores da chave criptográfica podem ter acesso ao conteúdo das informações, as quais ficam ininteligíveis àqueles que não possuem autorização para acessá-las[115].

A criptografia ponta a ponta constitui a técnica criptográfica mais popular, sendo amplamente utilizada, por exemplo, em aplicativos de troca de mensagens, como o *Messenger,* o *Signal* e o *WhatsApp*[116]. Por meio deste mecanismo somente o emissor e o destinatário da mensagem ("as pontas da comunicação"[117]) têm acesso à chave de decriptação

---

[111] BRASIL. Supremo Tribunal Federal. Arguição de Descumprimento de Preceito Fundamental 403/SE. **Voto do Relator Min. Luiz Edson Fachin**. p. 54. Disponível em: <https://www.conjur.com.br/dl/fachin-suspensao-whatsapp-decisao.pdf>. Acesso em 24 set. 2021.

[112] Os algoritmos são simétricos quando o mesmo conjunto (ou seja, a mesma chave) é utilizado para encriptar e decriptar a informação codificada. Esta chave é chamada de *single key*, e pode ser acessada apenas pelas partes envolvidas no processo de criptografia. Os algoritmos assimétricos são aqueles em que há dois conjuntos distintos: um para encriptação, que constitui uma chave pública (*public key*), e outro de decriptação, que constitui a chave privada (*single key*). Vide: PAAR, Christof; PELZL, Jan. **Understanding cryptography**: a textbook for students and practitioners. Londres: Springer, 2010. p. 5.

[113] GRESHAM, Joshua. **Is encrypted data personal data under the GDPR?** 6 mar. 2019. Disponível em: <https://iapp.org/news/a/is-encrypted-data-personal-data-under-the-gdpr/>. Acesso em 21 out. 2021.

[114] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 19. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

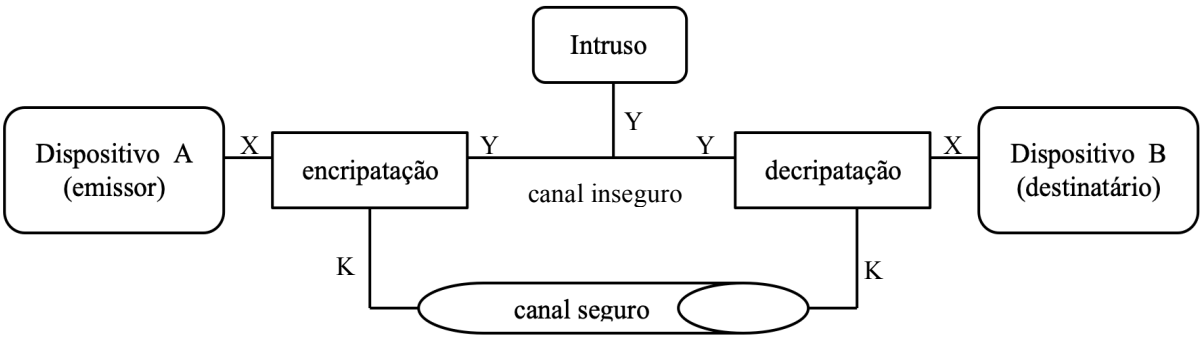
[115] SPINDLER, Gerald; SCHMECHEL, Philipp. Personal Data and Encryption in the European General Data Protection Regulation. **Journal of Intellectual Property, Information Technology and Electronic Commerce Law**, v. 7, n. 2, p. 163-177, set. 2016. p. 169.

[116] ARANHA, Diego F. O que é criptografia fim a fim e o que devemos fazer a respeito? In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-1.3.

[117] A depender da aplicação e da finalidade do uso da criptografia ponta a ponta, as pontas da comunicação podem ter diferentes naturezas, podendo ser pessoas naturais ou, até mesmo, sistemas automatizados. É o que acontece, por exemplo, nos protocolos SSL/TLS, que visam oferecer ao usuário uma segura navegação na rede, entre seu terminal e o servidor em que está hospedado certo sítio eletrônico acessado. Vide: ARANHA, Diego F. O que é criptografia fim a fim e o que devemos fazer a respeito? In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-1.3. e DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego

e, consequentemente, ao conteúdo daquela, denominado *plaintext*[118]. Cada mensagem enviada possui uma chave própria, de modo que se uma chave for decifrada, ter-se-á acesso ao conteúdo de apenas uma mensagem e não da totalidade da conversa[119]. Na imagem abaixo, o dispositivo A envia uma mensagem ao dispositivo B que é criptografada pela chave pública X, à qual tem acesso ambos os Dispositivos bem como o servidor do aplicativo. A mensagem criptografada pela chave pública passa por decriptação através da chave privada K, de posse do emissor e do destinatário da comunicação. O canal inseguro Y corresponde à mensagem criptografada ou *ciphertext*, de modo que pontos intermediários ou "invasores" apenas terão acesso a ela. O *plaintext* pode ser acessado apenas pelo Dispositivo A e pelo Dispositivo B, detentores da chave de decriptação K, mantendo-se confidencial o conteúdo da mensagem enviada[120].

Figura 1 – Sistema de Criptografia ponta a ponta.



Fonte: Adaptado a partir de PAAR, Christof; PELZL, Jan. **Understanding cryptography**: a textbook for students and practitioners. Londres: Springer, 2010. p. 5.

Como se pode observar acima, o acesso à chave privada é de suma importância para acessar o conteúdo dos dados criptografados e, consequentemente, para identificar eventuais sujeitos aos quais os dados se referem. Portanto, se o controlador mantiver a chave

---

(coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.3.

[118] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.3.

[119] WHATSAPP. **Sobre a criptografia de ponta a ponta.** Disponível em: <https://faq.whatsapp.com/general/security-and-privacy/end-to-end-encryption/?lang=pt_br> Acesso em 21 out. 2021.

[120] PAAR, Christof; PELZL, Jan. **Understanding cryptography**: a textbook for students and practitioners. Londres: Springer, 2010. p. 5.

criptográfica, a reidentificação dos titulares será plenamente possível. Neste caso, a criptografia configura uma técnica de pseudonimização, posto que a chave criptográfica poderá ser considerada a "informação adicional mantida separadamente pelo controlador em ambiente controlado e seguro" que permite a associação, direta ou indireta, de um dado a um indivíduo[121]. O dado criptografado permanece sendo categorizado como 'pessoal', sendo, pois, aplicáveis as regras da LGPD. De outro lado, no caso da criptografia de sentido único[122], a chave não é mantida pelo controlador, de modo que o procedimento, em tese, cria dados anonimizados[123]. Em sendo esta a situação, será necessária a análise da possibilidade de decriptação dos dados por terceiros, por meio do uso de meios razoáveis, para determinar se os dados criptografados encontram-se ou não fora do escopo material da LGPD. Se os elementos caracterizadores dos dados anonimizados não forem atendidos, a informação será qualificada como pessoal e o regime de proteção de dados é aplicável[124].

Danilo Doneda e Diego Machado citam três situações-tipo que bem ilustram como se pode definir se o regime de proteção de dados pessoais abrange ou não os dados criptografados[125]. A primeira delas é referente à Figura 1, qual seja, "a comunicação intermediada por provedor de aplicação de internet, em que o teor das mensagens e dados só pode ser acessado pelo(s) usuário(s) que possui(em) a chave criptográfica pertinente". Tanto em serviços de correio eletrônico quanto em aplicativos de mensagem instantânea, o conteúdo das mensagens em si não é sempre considerado dado pessoal, pois pode não ser relativo a uma pessoa. Todavia, ainda que o seja, a chave criptográfica privada é acessada apenas pelo emissor e pelo receptor da mensagem. O provedor do serviço tem acesso apenas à chave pública de encriptação, de modo que, para ele, os dados estão, a princípio, anonimizados[126]. A

---

[121] SPINDLER, Gerald; SCHMECHEL, Philipp. Personal Data and Encryption in the European General Data Protection Regulation. **Journal of Intellectual Property, Information Technology and Electronic Commerce Law**, v. 7, n. 2, p. 163-177, set. 2016. p. 169.

[122] A criptografia em sentido único ou função *hash* unidirecional é o termo utilizado na ciência da computação para para se referir a uma função matemática que produz um número sequencial ou um número gerado aleatoriamente que não é matematicamente derivado dos dados originais. Vide: EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 21. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.; e MERKLE, Ralph C. A fast software one-way hash function. **Journal of Cryptology**, v. 3, n. 1, p. 43-58, jul. 1990.

[123] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 18. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

[124] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.3.

[125] *Ibidem*.

[126] Não se está a considerar os dados fornecidos pelo usuário na criação da conta no aplicativo, porquanto estes dados, em geral, não são criptografados, e se tratam, indiscutivelmente, de dados pessoais.

segunda situação-tipo refere-se aos "dados de usuários armazenados e processados em servidores ou bases de dados de ente responsável, ou terceiro a ele interligado, encriptados por iniciativa ou determinação do próprio provedor de serviço, que pode acessar a chave de decriptação". O acesso à chave, como já abordado, permite a decodificação dos dados criptografados, e, por se tratar de dados de usuários, seria possível identificar os titulares. Neste caso, a criptografia configura pseudonimização, aplicando-se as regras da LGPD. Por fim, o terceiro cenário é concernente às "informações cifradas por ato do usuário e armazenadas e processadas em servidores ou bases de dados de provedor de serviço de computação em nuvem, o qual não tem acesso à correspondente chave criptográfica". Este mecanismo pode ser adotado em aplicativos como o Dropbox e o Google Drive, os quais já aplicam, por *default*, a criptografia, mas também abrem a possibilidade de que o próprio usuário, antes de enviar os arquivos para a "nuvem", realize, ele mesmo, a encriptação destes. Os dados serão duplamente encriptados: a primeira vez pelo cliente, único possuidor da chave privada desta encriptação; e a segunda vez pelo provedor do serviço, caso em que o usuário e o provedor tem acesso à chave criptográfica[127]. Como o provedor não têm acesso à chave da primeira encriptação, as informações criptografadas não podem ser por ele decifradas, permanecendo ininteligíveis e, por conseguinte, anonimizadas.

Nas ocasiões em que os dados criptografados podem ser considerados anonimizados (primeira e terceira situações-tipo), se faz necessário passar por outro plano de análise: avaliar se os meios utilizados para decifrar os dados e identificar os respectivos titulares podem ser razoavelmente utilizados pelo controlador e/ou por terceiros. Para tanto, deve-se avaliar três fatores objetivos: (i) a força do algoritmo de encriptação utilizado; (ii) a extensão (em bits) da chave de encriptação; (iii) se a chave está mantida em local seguro[128]. Tem-se, dessa forma, que não é correta a afirmação de que a criptografia configura técnica de anonimização de dados pelo simples fato de ser apta a tornar informações ininteligíveis[129]. Deve-se analisar a

---

[127] GOOGLE. **Primeiros passos com arquivos criptografados no Drive, no Documentos, no Planilhas e no Apresentações.** Disponível em: <https://support.google.com/docs/answer/10519333?hl=pt-BR>. Acesso em 21 out. 2021: "Qualquer usuário com quem um arquivo criptografado foi compartilhado pode acessar o arquivo com uma chave de criptografia exclusiva. Normalmente o Google criptografa o conteúdo em trânsito e em repouso, mas com a criptografia do lado do cliente, seu domínio decidiu adicionar mais uma camada de proteção. [...] A criptografia oferece uma camada extra de proteção para sua organização. Seus arquivos são criptografados de ponta a ponta e entre clientes. O Google não pode descriptografar os arquivos.".

[128] SPINDLER, Gerald; SCHMECHEL, Philipp. Personal Data and Encryption in the European General Data Protection Regulation. **Journal of Intellectual Property, Information Technology and Electronic Commerce Law**, v. 7, n. 2, p. 163-177, set. 2016. p. 172.

[129] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.3.

maneira pela qual o procedimento foi implementado, bem como quem tem acesso à chave de decriptação para, assente nisso, determinar se foi realizada a anonimização ou a pseudonimização dos dados.

3. A ANONIMIZAÇÃO DE DADOS PESSOAIS

A LGPD, assim como a maior parte das regulações de proteção de dados, adota uma lógica binária que divide os dados em duas grandes categorias: dados pessoais e dados não pessoais. Os dados que não são qualificados como pessoais são, consoante já referido, denominados anônimos ou anonimizados. Neste grupo, há informações que nunca foram vinculadas a uma pessoa natural identificada ou identificável e, portanto, nunca foram categorizadas como pessoais, e outras que já o foram, mas o vínculo fora removido[130]. A quebra da conexão antes estabelecida com o titular é realizada por meio de um processo denominado anonimização de dados, o qual é definido no inciso XI, do artigo 5º da LGPD como a "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". O dado perde, pois, a capacidade de revelar aspectos relativos ao sujeito que antes o titulava[131]. Em tese, o processo de anonimização é irreversível, sendo inviável a reidentificação do titular do dado, seja pelo controlador e pelo operador, seja por terceiros[132].

A anonimização consiste em um processo que é normalmente aplicado após a coleta dos dados pessoais. Logo, o pressuposto inicial é que os dados pessoais devem ter sido coletados em conformidade com a Lei, observando-se as bases legais que permitem o seu tratamento (consentimento, legítimo interesse, etc.) para que, depois, se proceda à anonimização destes[133]. Além disso, entre as operações, previstas na LGPD, que configuram "tratamento", está a "modificação". Tendo em vista que qualquer forma de anonimização envolverá, invariavelmente, a alteração dos dados, pode-se dizer que tal processo configura uma modalidade de tratamento[134]. Em seu art. 6º, inciso I, a LGPD prevê que para cumprir com o princípio da finalidade, a realização do tratamento deverá ter propósitos legítimos, específicos, explícitos e informados ao titular, os quais deverão ser compatíveis ao tratamento

---

[130] FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020. p. 13.
[131] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2 ed. São Paulo: Thomson Reuters Brasil, 2019. p. 140.
[132] NALDI, Maurizio. **Anonymization Systems and Utility**. IPEN - Workshop Rome, 12 jun. 2019. Disponível em: <https://edps.europa.eu/sites/default/files/publication/12-06-19_maurizio-naldi_anonymization-systems-and-utility_en_0.pdf>. Acesso em 22 out. 2021.
[133] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.
[134] EL EMAM, Khaled; HINTZE, Mike. **Does anonymization or de-identification require consent under the GDPR?**. 29 jan. 2019. Disponível em: <https://iapp.org/news/a/does-anonymization-or-de-identification-require-consent-under-the-gdpr/>. Acesso em 20 out. 2021.

posterior dos dados[135]. Assim, até o momento anterior de os dados pessoais se tornarem anonimizados, o seu tratamento deverá observar as regras dispostas em Lei, o que inclui a compatibilidade com a finalidade inicial pretendida com a coleta[136].

A Lei Brasileira pontua algumas modalidades de tratamento de dados pessoais que devem assegurar, sempre que possível, a anonimização. O uso de dados para estudos por órgãos de pesquisa é uma destas situações, o qual está previsto no artigo 7º, inciso IV[137], da LGPD. Em maio de 2020, no julgamento da ADIN 6389, a decisão histórica proferida pelo STF que reconheceu o direito fundamental à proteção de dados pessoais girou muito em torno da não observância das medidas de mitigação de riscos quando do tratamento de dados para fins de estatística. Na oportunidade, o Tribunal determinou a suspensão da eficácia da Medida Provisória 954/2020 (LGL\2020\4849), a qual, em seu art. 2º, *caput*, determinava "o compartilhamento de dados não anonimizados de telefonia fixa e móvel e o endereço de todos os brasileiros com o IBGE – Fundação Instituto Brasileiro de Geografia e Estatística"[138]. O acesso a estes dados teria como objetivo dar suporte à produção estatística oficial durante a situação de distanciamento social decorrente da pandemia do COVID-19[139]. Uma das questões controversas foi a ausência de implementação de medidas de segurança para que o compartilhamento pudesse se dar sem risco de acessos indevidos[140], o que seria, possivelmente, solucionado caso os dados tivessem passado por processo de anonimização. À época do julgamento, a LGPD ainda não havia entrado em vigor, razão pela qual a orientação de garantia da anonimização dos dados prevista no inciso IV do artigo 7º não poderia ser

---

[135] 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; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[136] EL EMAM, Khaled; HINTZE, Mike. **Does anonymization or de-identification require consent under the GDPR?**. 29 jan. 2019. Disponível em: <https://iapp.org/news/a/does-anonymization-or-de-identification-require-consent-under-the-gdpr/>. Acesso em 20 out. 2021.

[137] Art. 7º. O tratamento de dados pessoais somente poderá ser realizado nas seguintes hipóteses: IV - para a realização de estudos por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais; BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[138] MENDES, Laura Schertel. **A encruzilhada da proteção de dados no Brasil e o caso do IBGE**. 23 abr. 2020. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/a-encruzilhada-da-protecao-de-dados-no-brasil-e-o-caso-do-ibge-23042020#sdfootnote2sym>. Acesso em 21 out. 2021.

[139] VITAL, Danilo. **Rosa Weber atende OAB e suspende MP do compartilhamento de dados**. 24 abr. 2020. Disponível em: < https://www.conjur.com.br/2020-abr-24/rosa-atende-oab-suspende-mp-compartilhamento-dados>. Acesso em 20 set. 2021.

[140] MENDES, Laura Schertel. **A encruzilhada da proteção de dados no Brasil e o caso do IBGE**. 23 abr. 2020. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/a-encruzilhada-da-protecao-de-dados-no-brasil-e-o-caso-do-ibge-23042020#sdfootnote2sym>. Acesso em 21 out. 2021.

exigida em sede de julgamento[141]. Contudo, a implementação deste procedimento foi, por diversas vezes, mencionada nos votos e sustentações orais, estando, inclusive, na Ementa da referida decisão:

> 6. Ao não apresentar mecanismo técnico ou administrativo apto a proteger, de acessos não autorizados, vazamentos acidentais ou utilização indevida, seja na transmissão, seja no tratamento, o sigilo, a higidez e, quando o caso, o anonimato dos dados pessoais compartilhados, a MP nº 954/2020 descumpre as exigências que exsurgem do texto constitucional no tocante à efetiva proteção dos direitos fundamentais dos brasileiros.[142]

A recomendação de anonimizar, sempre que possível, os dados utilizados para fins de pesquisa se justifica pelo fato de que, neste contexto de processamento de dados, a identificação dos titulares é dispensável a depender do objetivo visado pelo estudo. Nesse sentido, a relatora Ministra Rosa Weber, em seu voto, destaca que "em momento algum a identificação dos indivíduos titulares dos dados foi reivindicada como necessária ao relevante trabalho desenvolvido pelo IBGE"[143]. Órgãos públicos e instituições privadas que pretendem desenvolver pesquisas devem verificar se o propósito visado pode ser alcançado sem que os titulares sejam identificados. Se isso for possível, a anonimização deve ser providenciada tanto como estratégia de diminuição de riscos, quanto como em cumprimento ao princípio da necessidade, segundo o qual o processamento de dados deve se limitar à abrangência dos dados pertinentes, proporcionais e não excessivos em relação às finalidades pretendidas[144]. Ao contrário do que se passou com a MP 954/2020, o programa Simi – Sistema de

---

[141] Consoante disposto no art. 1º da Lei de Introdução às Normas do Direito Brasileiro, a lei, salvo disposição contrário, "começa a vigorar em todo o País 45 (quarenta e cinco) dias depois de oficialmente publicada". BRASIL. **Decreto-lei nº 4.657, de 4 de setembro de 1942**. Lei de Introdução às Normas do Direito Brasileiro. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto-lei/del4657compilado.htm>. Acesso em 20 out. 2021.

[142] BRASIL. Supremo Tribunal Federal. ADIs nº 6.387, 6.388, 6.389, 6.390 e 6.393. Relatora: Ministra Rosa Weber. 7 de maio de 2020. p. 2. Disponível em: <http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895165> Acesso em 21 set. 2021.

[143] BRASIL. Supremo Tribunal Federal. ADIs nº 6.387, 6.388, 6.389, 6.390 e 6.393. Relatora: Ministra Rosa Weber. 7 de maio de 2020. p. 16. Disponível em: <http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895165> Acesso em 21 set. 2021.

[144] Art. 6º. As atividades de tratamento de dados pessoais deverão observar a boa-fé e os seguintes princípios: [...] III - necessidade: limitação do tratamento ao mínimo necessário para a realização de suas finalidades, com abrangência dos dados pertinentes, proporcionais e não excessivos em relação às finalidades do tratamento de dados; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

Monitoramento Inteligente, iniciativa governamental do estado de São Paulo instituída pelo Decreto Estadual 64.963/2020, que também tinha como objetivo o acesso a dados para a formulação de estratégias e políticas públicas de combate à pandemia de Covid-19[145], obteve resposta positiva do Poder Judiciário ao ser contestado na ação popular nº. 1020192-74.2020.8.26.0053, em junho de 2020[146]. A decisão que manteve o funcionamento da medida, baseou-se amplamente na manifestação da Procuradoria Geral do Estado de São Paulo que, fundada no Acordo de Cooperação Técnica, informou que os dados tratados pelo SIMI são estatísticos, agregados e anônimos, sem identificação dos clientes, em consonância com o art. 5º, III da LGPD[147].

O outro contexto em que a Lei indica a adoção da anonimização é no uso de dados para o desenvolvimento de pesquisas em saúde pública e/ou que lidem com dados pessoais sensíveis, segundo dispõe, respectivamente, os artigos 13[148] e 11, inciso II, alínea "c"[149]. A peculiaridade acima exposta relativa à dispensabilidade da identificação dos titulares quando o tratamento dos dados é realizado para fins de pesquisa soma-se ao fato de os dados em questão serem qualificados como sensíveis, o que constitui outro fundamento para a adoção da anonimização nestes casos. Os dados pessoais sensíveis, por estarem relacionados a aspectos da personalidade que tornam o titular vulnerável à discriminação (estado de saúde, orientação sexual, posicionamento político, dados genéticos, etc.) apresentam maiores riscos

---

[145] CORRÊA, Adriana Espíndola; DA LUZ, Pedro Henrique Machado. **A exceção na proteção de dados pessoais durante a Covid-19 - parte 1**. 22 mai. 2020. Disponível em: <https://www.conjur.com.br/2020-mai-22/direito-civil-atual-excecao-protecao-dados-pessoais-durante-covid-19>. Acesso em 24 out. 2021.

[146] SÃO PAULO. 3ª Vara de Fazenda Pública. Ação Popular 1020192-74.2020.8.26.0053. Requerente: Mauricio Roberto Giosa. Requerido: Fazenda Pública do Estado de São Paulo. Julgador: Juiz Luis Manuel Fonseca Pires. São Paulo, 22 jun. 2020. Disponível em: <https://esaj.tjsp.jus.br/cpopg/show.do?processo.codigo=1H000HDVE0000&processo.foro=53&processo.numero=1020192-74.2020.8.26.0053&uuidCaptcha=sajcaptcha_bfe33b253a5443ef808ae42740df88c3>. Acesso em 24 out. 2021.

[147] CORRÊA, Adriana Espíndola; DA LUZ, Pedro Henrique Machado. **A exceção na proteção de dados pessoais durante a Covid-19 – parte 2**. 23 mai. 2020. Disponível em: <https://www.conjur.com.br/2020-mai-23/direito-civil-atual-excecao-protecao-dados-pessoais-durante-covid-19>. Acesso em 24 out. 2021.

[148] Art. 13. Na realização de estudos em saúde pública, os órgãos de pesquisa poderão ter acesso a bases de dados pessoais, que serão tratados exclusivamente dentro do órgão e estritamente para a finalidade de realização de estudos e pesquisas e mantidos em ambiente controlado e seguro, conforme práticas de segurança previstas em regulamento específico e que incluam, sempre que possível, a anonimização ou pseudonimização dos dados, bem como considerem os devidos padrões éticos relacionados a estudos e pesquisas. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[149] Art. 11. O tratamento de dados pessoais sensíveis somente poderá ocorrer nas seguintes hipóteses: [...] II - sem fornecimento de consentimento do titular, nas hipóteses em que for indispensável para: [...] c) realização de estudos por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais sensíveis; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

potenciais de lesão àquele, se houver algum tipo de vazamento durante o processamento[150]. Por esta razão, a Lei confere especial grau de proteção a essa espécie de dados, de modo que, sempre que forem tratados, o titular deve ter consentido expressamente. O tratamento de dados sensíveis pode apenas ocorrer sem o consentimento do titular caso seja indispensável para a realização de estudos por órgãos de pesquisa, situação em que a anonimização deve ser garantida, sempre que possível. Durante a crise mundial ocasionada pela pandemia do COVID-19, os dados pessoais tornaram-se grandes aliados na execução de políticas de controle e de contenção do vírus[151]. Muitas das medidas adotadas, não só no Brasil, mas ao redor de todo o globo, lidam com dados relativos ao estado de saúde dos pacientes contaminados pelo vírus, os quais são, portanto, categorizados como dados sensíveis. Sistemas de *contact tracing*[152], por exemplo, "funcionam com a troca de identificadores anônimos entre telefones próximos via Bluetooth após a instalação de um aplicativo disponibilizado pela autoridade de saúde nacional ou eventualmente pelo próprio sistema operacional, a depender de como opera a solução"[153]. Este mecanismo visa mapear a rede de contatos que uma pessoa contaminada pelo vírus teve, o que depende, por óbvio, que aquela informe no aplicativo que o seu teste de COVID-19 obteve resultado positivo. Para garantir a proteção dos direitos e liberdades individuais das pessoas que forneceram tais dados, os identificadores são anonimizados, impedindo que se possa identificá-las.

A anonimização é também um direito garantido pela Lei ao titular de dados pessoais quando os dados forem desnecessários, excessivos ou tratados em desconformidade com a legislação[154], podendo ele requerer ao controlador, a qualquer momento, que o referido processo seja implementado. Trata-se de uma das maneiras pelas quais o titular exerce o seu direito à autodeterminação informativa, um dos fundamentos da disciplina da proteção de

---

[150] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019. p. 142-143.

[151] DONEDA, Danilo. **A proteção de dados em tempos de coronavírus**. 25 mar. 2020. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/a-protecao-de-dados-em-tempos-de-coronavirus-25032020>. Acesso em 21 out. 2021.

[152] De acordo com o Conselho da Europa, há aproximadamente 52 aplicativos destinados ao rastreamento de contatos (*contact tracing apps*) em funcionamento ao redor do mundo. Vide: COUNCIL OF EUROPE. **Contact Tracing Apps**. 10 jun. 2020. Disponível em: <https://www.coe.int/en/web/data-protection/contact-tracing-apps>. Acesso em 24 out. 2021.

[153] ALMEIDA, Bethania de Araujo et al. Preservação da privacidade no enfrentamento da COVID-19: dados pessoais e a pandemia global. **Ciência & Saúde Coletiva [online]**, v. 25, p. 2487-2492, jun. 2020. Disponível em: <https://doi.org/10.1590/1413-81232020256.1.11792020>. Acesso em 24 out. 2021.

[154] Art. 18. O titular dos dados pessoais tem direito a obter do controlador, em relação aos dados do titular por ele tratados, a qualquer momento e mediante requisição: [...] IV - anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade com o disposto nesta Lei; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

dados pessoais, que se traduz na ideia de que o indivíduo deve ter o poder de, ele próprio, decidir acerca da divulgação e utilização de seus dados pessoais[155].

Nesta mesma linha, o § 6º do artigo 18 da LGPD estabelece que o responsável pelo tratamento dos dados deverá informar, de maneira imediata, aos agentes de tratamento com os quais tenha realizado uso compartilhado de dados a correção, a eliminação, a anonimização ou o bloqueio dos dados, para que repitam idêntico procedimento, se o titular dos dados houver revogado o seu consentimento. A LGPD obriga, por via de regra, a eliminação dos dados pessoais quando finda a condição que autorizou o seu tratamento[156]. Ao revogar o consentimento, base legal que propiciou o processamento dos seus dados pessoais, o titular está comunicando o seu interesse no fim do tratamento, motivo pelo qual o controlador deverá adimplir aos requisitos de término estabelecidos em Lei. Nas ocasiões que envolvam a transferência dos dados a terceiro, o artigo 16, inciso III[157], autoriza a conservação dos dados pelo controlador após o final do tratamento. Por esta razão, abre-se a possibilidade de os dados passarem por processo de anonimização ao invés de serem eliminados. A eliminação após o término do "ciclo de vida" dos dados pessoais ressalva, também, a hipótese em que o controlador fará o uso exclusivo dos dados, sem possibilitar o acesso a terceiros[158]. Em tal situação, a anonimização deverá ser, obrigatoriamente, implementada.

3.1 A REVERSIBILIDADE DO PROCESSO DE ANONIMIZAÇÃO

Os conceitos normativos adotados pelas regulações de proteção de dados pessoais derivam das primeiras gerações de leis[159], sendo, pois, cunhados quando se tinha um outro

---

[155] MENKE, Fabiano. **As origens alemãs e o significado da autodeterminação informativa**. 30 out. 2020. Disponível em: <https://www.migalhas.com.br/coluna/migalhas-de-protecao-de-dados/335735/as-origens-alemas-e-o-significado-da-autodeterminacao-informativa>. Acesso em 24 out. 2021.

[156] JR., Sérgio Alves. Fechando um ciclo: do término do tratamento de dados pessoais (arts. 15 e 16 da LGPD). In: DONEDA, Danilo et al (coord.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. p. 227-241.

[157] Art. 16. Os dados pessoais serão eliminados após o término de seu tratamento, no âmbito e nos limites técnicos das atividades, autorizada a conservação para as seguintes finalidades: [...] II - transferência a terceiro, desde que respeitados os requisitos de tratamento de dados dispostos nesta Lei; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[158] Art. 16. Os dados pessoais serão eliminados após o término de seu tratamento, no âmbito e nos limites técnicos das atividades, autorizada a conservação para as seguintes finalidades: [...] IV - uso exclusivo do controlador, vedado seu acesso por terceiro, e desde que anonimizados os dados. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[159] Viktor Mayer-Schönberger propõe a divisão dos regulamentos de proteção de dados pessoais em quatro diferentes gerações, baseadas nas diferenças e similitudes dos regimes adotados ao longo dos anos, especialmente na Europa. A primeira geração de leis corresponde àquelas regulações elaboradas entre as décadas de 60 e 70, tendo sido as pioneiras a tentar oferecer uma resposta ao surgimento do processamento eletrônico de

estado da arte da tecnologia. Hoje, com o avanço tecnológico e a disponibilização massiva de dados das mais variadas fontes, o cenário da proteção de dados foi alterado. A diferenciação entre dados pessoais e dados não pessoais, ou seja, anônimos, encontra-se cada vez mais dificultada na prática. Ao passo que as técnicas de anonimização de dados se tornam cada vez mais sofisticadas, as contra-tecnologias também evoluem na mesma proporção.

Estudos realizados no campo da ciência da computação demonstram que processos de anonimização, antes tidos como confiáveis, apresentam falhas, podendo ser revertidos[160]. Um dos precursores destes trabalhos foi desenvolvido no ano de 2000 por Latanya Sweeney[161]. Seu propósito era determinar quantos indivíduos poderiam ser identificados a partir dos dados disponibilizados pelo Censo dos Estados Unidos da América de 1990. Com base em apenas três tipos diferentes de informação (código postal, gênero, data de nascimento), constatou-se que 87% (216 milhões de 248 milhões) da população dos Estados Unidos à época poderia ser identificada de forma única. Sweeney demonstrou que a retirada de identificadores diretos, denominados por ela de "identificadores explícitos"[162], como o nome e o número de celular, não são suficientes para considerar que uma base de dados foi anonimizada[163], já que, por meio da combinação de outros atributos (combinações únicas), é possível identificar o titular dos dados. Ao conjunto de atributos que formam combinações únicas, ou quase únicas, capazes de identificar um indivíduo, Sweeney dá o nome de *quasi-identifier,* no português, "quase-identificador"[164]. A combinação formada por código postal, gênero e data de nascimento é, consoante os resultados do estudo, um quase-identificador único para grande parte da população estadunidense. O método utilizado pela pesquisadora para chegar a tal

---

dados dentro do Governo e das grandes corporações. Vide: MAYER-SCHÖNBERGER Viktor. Generational Development of Data Protection in Europe. In: AGRE, Philip E.; ROTENBERG, Marc (orgs.). **Technology and Privacy:** The New Landscape. Cambridge: The MIT Press, 1997, p. 219-241.

[160] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.2.

[161] SWEENEY, Latanya. Simple demographics often identify people uniquely. **Health (San Francisco)**, v. 671, n. 2000, p. 1-34, 2000.

[162] *Ibidem,* p. 6: "An explicit identifier is a set of data elements, such as {name, address} or {name, phone number}, for which there exists a direct communication method, such as email, telephone, postal mail, etc., where with no additional information, the designated person could be directly and uniquely contacted.".

[163] Às bases de dados que tiveram apenas os identificadores diretos retirados, Sweeney dá o nome de *de-identified data*, que, em tradução literal, chamar-se-ia de dados "desidentificados". Vide: SWEENEY, Latanya. Simple demographics often identify people uniquely. **Health (San Francisco)**, v. 671, n. 2000, p. 1-34, 2000. p. 7: "De-identified data result when all explicit identifiers, such as name, address, or phone number are removed, generalized or replaced with a made-up alternative.".

[164] *Ibidem*:"A quasi-identifier is a set of data elements in entity-specific data that in combination associates uniquely or almost uniquely to an entity and therefore can serve as a means of directly or indirectly recognizing the specific entity that is the subject of the data.".

conclusão foi a correlação dos dados fornecidos no Censo, os quais estavam publicados na *Web*, com dados coletados em hospitais, clínicas médicas e laboratórios, localizados em diferentes estados do país. Estes últimos foram disponibilizados pela *National Association of Health Data Organizations* – NAHDO a pesquisadores, profissionais da indústria e, em alguns estados, ao público em geral.

Outro estudo de suma relevância foi realizado em 2007 por Arvind Narayanan e Vitaly Shmatikov. Em 2006, a Netflix, atual líder global no mercado de *streaming* que, em 2019, atingiu US$ 20 bilhões de receita anual[165], anunciou um prêmio no valor de U$ 1 mihão para o ganhador de um concurso promovido pela empresa que visava melhorar os seus serviços de recomendação de filmes[166]. Para tanto, a Netflix divulgou um conjunto de dados contendo 100.480.507 classificações de filmes, realizadas por 480.189 assinantes entre os anos de 1999 e 2005. De acordo com a empresa, os dados eram anonimizados, porquanto toda a informação que possibilitava a identificação de cada cliente havia sido removida, mantendo-se somente a classificação do filme e a respectiva data, dados esses que haviam sido submetidos à randomização[167]. Com base nestas informações, Narayanan e Shmatikov desenvolveram um estudo para testar a anonimização dos dados implementada pela Netflix[168]. O objetivo dos pesquisadores era verificar quanto de conhecimento sobre um assinante da Netflix um adversário[169] precisaria ter para identificar o registro daquele no conjunto de dados anonimizados disponibilizados pela empresa, e, assim, acessar o seu histórico completo de visualização de filmes. A resposta encontrada foi: muito pouco[170]. Consoante os resultados da pesquisa, supondo-se que o adversário tivesse acesso a algumas classificações concedidas por determinado assinante e às datas correspondentes, seriam necessárias oito classificações (dentre as quais duas poderiam estar completamente erradas) e datas que poderiam ter uma

---

[165] Netflix annual revenue hits US$20 billion. 1 vídeo (4 min). Publicado por BNN – Bloomberg. 22 jan. 2020. Disponível em: <https://www.bnnbloomberg.ca/technology/video/netflix-annual-revenue-hits-us-20-billion~1881767>. Acesso em 25 out. 2021.

[166] NARAYANAN, Arvind; SHMATIKOV, Vitaly. How To Break Anonymity of the Netflix Prize Dataset. **ArXiv e-prints**, out. 2006. Disponível em: <https://arxiv.org/abs/cs/0610105#>. Acesso em 25 out. 2021.

[167] *Ibidem*. p. 1-2.

[168] *Ibidem*. p. 2.

[169] "Adversário" é o termo adotado pela literatura técnica para aquele que pretende reverter o processso de anonimização. Vide: OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. **UCLA Law Review**, v. 57, n. 6, p. 1701-1778, 2010. p. 1707-1708: "A person, known in the scientific literature as an adversary, reidentifies anonymized data by linking anonymized records to outside information, hoping to discover the true identity of the data subjects.".

[170] NARAYANAN, Arvind; SHMATIKOV, Vitaly. How To Break Anonymity of the Netflix Prize Dataset. **ArXiv e-prints**, out. 2006. Disponível em: <https://arxiv.org/abs/cs/0610105#>. Acesso em 25 out. 2021. p. 2: "Anonymity of the Netflix dataset thus depends on the answer to the following question: How much does the attacker need to know about a Netflix subscriber in order to identify her record in the dataset, and thus learn her complete movie viewing history? In the rest of this paper, we investigate this question. In brief, the answer is: very little.*"*

margem de erro de três dias, para que 96% dos assinantes, cujos registros foram disponibilizados, fossem identificados de forma inequívoca no conjunto de dados[171]. Para chegar a esta conclusão, os pesquisadores correlacionaram informações publicamente disponíveis no *Internet Movie Database* – IMDb[172] àquelas constantes no conjunto de dados disponibilizados pela Netflix. Outro ponto importante do estudo é que, por meio desta combinação com as informações publicadas no IMDb, Narayanan e Schmatikov constataram ser possível acessar dados sensíveis e não públicos dos assinantes, como, por exemplo, suas preferências políticas ou mesmo orientação sexual[173].

Com o crescente número de publicações que comprovam falhas em diversas técnicas de anonimização de dados, o entendimento no sentido de que a anonimização poderia ser absoluta foi sendo abandonado. Durante décadas, acreditou-se que seria possível proteger eficientemente a privacidade das pessoas por meio de pequenas alterações nos seus dados[174], como, por exemplo, a eliminação de identificadores diretos/explícitos, consoante realizado no Censo dos EUA de 1990 e nos dados divulgados pela Netflix. Paul Ohm chama esta antiga "crença" de suposição da anonimização robusta[175] (*robust anonymisation assumption*)[176]. Até a década de 1990, a eficácia da anonimização não era questionada, o que funcionava bem para todos os envolvidos: os responsáveis pelo tratamento de dados alegavam proteger a privacidade dos titulares ao compartilhar os dados com terceiros; as pessoas naturais em causa confiavam que seus dados estavam seguros; os legisladores acreditavam poder equilibrar a privacidade e outros interesses (tais como o avanço do conhecimento) por meio da não aplicação de regras de privacidade e proteção de dados ao tratamente de dados anonimizados; e os reguladores podiam facilmente dividir os agentes de tratamento de dados em responsáveis (aqueles que anonimizavam) e irresponsáveis (aqueles que não o faziam)[177].

As primeiras leis de proteção de dados foram elaboradas com base nesta (hoje, superada) ideia de que a anonimização consiste em um processo 100% eficaz. Quando a sua

---

[171] *Ibidem.*
[172] O IMDb é uma base de dados *online* na qual podem ser publicadas classificações e críticas sobre filmes e programas de televisão.
[173] NARAYANAN, Arvind; SHMATIKOV, Vitaly. How To Break Anonymity of the Netflix Prize Dataset. **ArXiv e-prints**, out. 2006. p. 3. Disponível em: <https://arxiv.org/abs/cs/0610105#>. Acesso em 25 out. 2021.
[174] OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. **UCLA Law Review**, v. 57, n. 6, p. 1701-1778, 2010. p. 1706-1707.
[175] Tradução adotada por Doneda e Machado em DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e criptografia: tecnologias criptográficas entre anonimização e pseudonimização de dados. IN: DONEDA, Danilo; MACHADO, Diego. coord. **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book.* p. RB-8.3.
[176] OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. **UCLA Law Review**, v. 57, n. 6, p. 1701-1778, 2010. p. 1706-1707.
[177] *Ibidem*. p. 1716.

irreversibilidade começou a ser contestada, várias regulações passaram a lidar com o seguinte problema: aquelas que admitiam ser meramente identificável a pessoa natural à qual o dado categorizado como pessoal se refere, e, ao mesmo tempo, estabeleciam uma dicotomia entre dados anonimizados e dados pessoais, apresentavam uma redundância normativa[178]. Isso porque, após o reconhecimento da possibilidade de reidentificação do titular por meio da reversão do processo de anonimização, dados anonimizados, assim como dados pessoais, tornariam a pessoa natural titular de dados identificável. Logo, o conceito de dados pessoais estender-se-ia aos dados anonimizados, os quais estariam, portanto, dentro do escopo de aplicação das leis de proteção de dados. A divisão da informação em pessoal e não pessoal (anonimizada) seria, dessa forma, uma falsa dicotomia[179].

O grande desafio gira em torno das chamadas *personally identifiable information* (PII), isso é, informações pessoais identificáveis. Ohm refere que os estudos acima relatados demonstraram que toda informação poderia ser classificada como PII para aquele que tem acesso à informação externa "certeira"[180]. O difícil, no entanto, é prever o tipo e a quantidade de informações auxiliares que o adversário poderia ter acesso[181]. Com o crescente papel que as redes sociais ocupam na vida das pessoas, a obtenção de informações públicas que possibilitam reidentificar uma base de dados é cada vez mais facilitada. O fenômeno da agregação de informações, que consiste na combinação de vários fragmentos de dados, se torna possível graças à considerável quantidade de dados que é diariamente disponibilizada, muitas vezes, pelos próprios titulares. Neste cenário, quanto mais informações se tem sobre determinada pessoa, maior a probabilidade de estas informações serem utilizadas para identificá-la em uma base de dados anonimizados[182]. Por exemplo, suponha-se que os dados da coluna A são considerados anonimizados por determinada empresa de marketing. A coluna B refere-se aos dados compartilhados pelo próprio indivíduo em seu perfil público de determinada rede social, e a coluna C aos dados fornecidos por ele a uma loja de *e-commerce*. Ao combinar as colunas, propicia-se a reidentificação do titular dos dados da coluna A, qual seja Carlos Silva.

---

[178] BIONI, Bruno. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo et al (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 40.
[179] NARAYANAN, Arvind; SHMATIKOV, Vitaly. How To Break Anonymity of the Netflix Prize Dataset. **ArXiv e-prints**, out. 2006. p. 4. Disponível em: <https://arxiv.org/abs/cs/0610105#>. Acesso em 25 out. 2021.
[180] OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. **UCLA Law Review**, v. 57, n. 6, p. 1701-1778, 2010. p. 1723: "These results suggest that maybe everything is PII to one who has access to the right outside information.".
[181] *Ibidem*. p. 1724.
[182] SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. **New York University Law Review**, v. 86, n. 6, p. 1814-1894, 2011. p. 1843.

Quadro 3 – Agregação de informações provenientes de fontes diversas.

COLUNA A	COLUNA B	COLUNA C
NOME: [redacted] **IDADE:** 16 **CEP:** 04849-529 **ESPORTE PREFERIDO:** Skate **HOBBIE:** Ouvir música **CURSO:** Administração	**NOME:** Caco Silva **IDADE:** 16 **LOCALIDADE:** São Paulo, BR **BIOGRAFIA:** Apaixonado por música e skatista nas horas vagas. Cursando Adm na USP.	**NOME:** Carlos Silva **DATA DE NASCIMENTO:** 27.07.2005 **ENDEREÇO:** Rua 13 de Maio, n. 105, apartamento 1409 **CEP:** 04849-529 **CIDADE:** São Paulo **ESTADO:** São Paulo

Fonte: Adaptado a partir de SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. **New York University Law Review**, v. 86, n. 6, p. 1814-1894, 2011. p. 1844.

Os dados da coluna A estariam anonimizados e sem risco de reversão caso não fosse possível o acesso a informações auxiliares provenientes de outras fontes. Contudo, considerando as circunstâncias próprias da sociedade da informação, seria equivocado assumir que o adversário jamais seria capaz de acessar outras informações para agregá-las àquelas constantes na base de dados anonimizados e, assim, reidentificar o titular. Na literatura técnica da segurança da informação, esta atitude desacreditada é chamada de "segurança por meio da obscuridade" (no inglês, *security through obscurity* – STO), segundo a qual é criada uma falsa ideia de segurança e observação às práticas de privacidade baseada na crença em que se as vulnerabilidades forem escondidas, então o sistema é seguro[183]. Para evitar esse tipo de visão deturpada da realidade, os especialistas assumem que o adversário sempre poderá possuir os fragmentos exatos de dados necessários para reidentificar titulares de informações anonimizadas, a fim de conceber respostas que protejam os dados, mesmo na pior conjuntura[184]. Partindo do pressuposto de que profissionais técnicos que implementam a anonimização de dados sempre reconhecem o risco de os titulares serem reidentificados, o Direito também precisou se adaptar à essa conjuntura e não mais se embasar na suposição da anonimização robusta.

A maioria das atuais leis de proteção de dados, incluindo a LGPD e o GDPR, adotam uma abordagem baseada no risco (*risk based approach*), ou seja, reconhecem que qualquer

[183] OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. **UCLA Law Review**, v. 57, n. 6, p. 1701-1778, 2010. p. 1724.
[184] *Ibidem.*

dado anonimizado detém o risco residual inerente de se transmudar em um dado pessoal[185]. O entendimento majoritário é no sentido de que o processo de anonimização gerencia circunstancialmente a identificabilidade de uma base de dados, de modo que para determinar se a reidentificação foi suficientemente afastada, deve-se realizar uma análise contextual[186]. O que significa dizer, utilizando-se as palavras de Bruno Bioni, que:

> [...] não há um único método ou uma combinação perfeita *ex ante* para parametrizar o processo de anonimização, devendo-se analisar contextualmente como este deve ser empreendido para que os titulares dos dados anonimizados não sejam reidentificados, nem mesmo por quem procedeu à sua anonimização.[187]

A depender do tipo de dado pessoal, o regime de proteção demandado pode variar: alguns devem ser mantidos confidenciais por disposição legal (por exemplo, dados estatísticos colhidos para o Censo Demográfico[188]), outros exigem maior atenção por parte do controlador e do operador por tornarem o titular suscetível à discriminação (dados pessoais sensíveis no geral), e assim por diante[189]. Da mesma forma, com base nas especificidades de cada dado e na finalidade do tratamento realizado, determina-se a(s) técnica(s) de anonimização mais efetiva(s) e apropriada(s). Por exemplo, em algumas bases de dados a aplicação de apenas uma técnica de anonimização é suficiente para afastar a possibilidade de reidentificar o titular, em outras, todavia, necessária a aplicação de uma combinação de diferentes técnicas de anonimização para garantir o afastamento da identificabilidade do titular. Além disso, é fundamental conhecer os principais pontos fortes e fracos de cada técnica para melhor escolher o processo de anonimização adequado a um dado contexto[190].

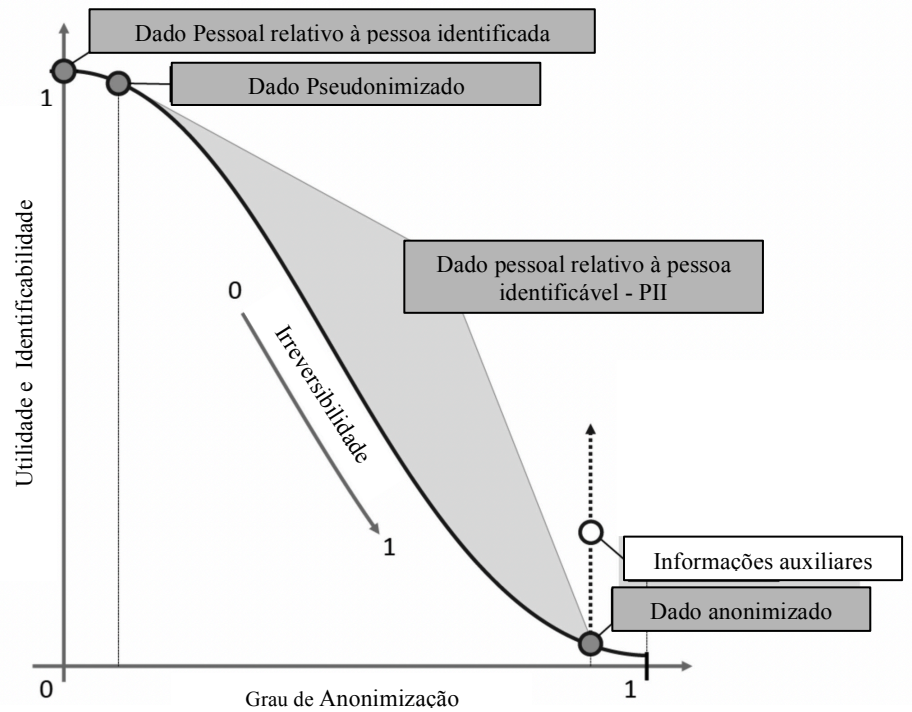
---

[185] BIONI, Bruno Ricardo. **Proteção de dados pessoais**: a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019. p. 75.
[186] *Ibidem*. p. 72.
[187] *Ibidem.*
[188] Art. 1º [...] Parágrafo único. As informações prestadas terão caráter sigiloso, serão usadas exclusivamente para fins estatísticos, e não poderão ser objeto de certidão, nem, em hipótese alguma, servirão de prova em processo administrativo, fiscal ou judicial, excetuado, apenas, no que resultar de infração a dispositivos desta lei. BRASIL. **Lei nº 5.534, de 14 de novembro de 1968**. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l5534.htm>. Acesso em 20 out. 2021.
[189] DOLIN, Ron A. Search Query Privacy: The Problem of Anonymization. **Hastings Science and Technology Law Journal**, v. 2, n. 2, p. 137-182, 2010. p. 140.
[190] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

3.2 TÉCNICAS DE ANONIMIZAÇÃO DE DADOS E SUAS FRAGILIDADES

Os dados apenas tem valor pois podem ser processados e, assim, permitem a obtenção de benefícios e o acesso a conhecimento. Se a retirada de identificadores for feita ao ponto de tornar os dados sem utilidade, a razão de ser do tratamento de dados deixa de existir. Deve-se atentar, portanto, que a utilidade de determinado dado é diretamente proporcional ao seu grau de identificabilidade[191]. Em outras palavras, quanto mais anonimizado estiver um dado, mais a sua utilidade decai. O gráfico abaixo (Figura 2) retrata esta relação de proporcionalidade: dados relativos a pessoas identificadas e dados pseudonimizados detêm os maiores níveis de utilidade e alto grau de identificabilidade; ao passo que o grau de identificabilidade e, consequentemente, a possibilidade de reversão do processo de anonimização decaem, a utilidade dos dados diminui[192]. Ainda, se aos dados anonimizados forem agregadas informações auxiliares, o grau de identificabilidade aumenta, bem como a sua utilidade. Por conta disso, a partir da análise caso a caso, deve-se tentar estabelecer um equilíbrio entre o grau de identificabilidade e a utilidade dos dados, sem que isso impossibilite alcançar a finalidade pretendida por meio do tratamento.

Figura 2 – Utilidade e Identificabilidade X Grau de Anonimização.



[191] VOKINGER, Kerstin N., et al. Lost in Anonymization — A Data Anonymization Reference Classification Merging Legal and Technical Considerations. **The Journal of Law, Medicine & Ethics**, v. 48, n. 1, p. 228–231, 2020. p. 230.
[192] *Ibidem*.

Fonte: Adaptado a partir de VOKINGER, Kerstin N., et al. Lost in Anonymization — A Data Anonymization Reference Classification Merging Legal and Technical Considerations. **The Journal of Law, Medicine & Ethics**, v. 48, n. 1, p. 228–231, 2020. p. 230.

Apesar de, atualmente, os limites técnicos e teóricos da anonimização de dados serem reconhecidos, isso não quer dizer que este processo sempre falha em proteger a privacidade e os dados das pessoas naturais em questão[193]. Existem diversas técnicas de anonimização que podem fornecer considerável nível de proteção ao titular de dados, mas, para isso, a sua aplicação deve ser corretamente gerenciada para se ajustar ao contexto e aos objetivos da atividade de tratamento dos dados. Estas técnicas variam, especialmente, em custo, complexidade, facilidade de implementação e robustez[194]. Não há na LGPD qualquer indicação de uma técnica de anonimização padrão, sendo concedida a liberdade de determinar a mais apropriada àquele que implementará o processo.

O Parecer 05/2014, elaborado pelo WP29, estabelece três critérios para avaliar a solidez de determinada técnica de anonimização de dados: **(i)** *singling out* ou distinção - possibilidade de isolar alguns ou todos os registros[195] que identificam um indivíduo em um conjunto de dados; **(ii)** *linkability* ou possibilidade de ligação - capacidade de ligar, pelo menos, dois registros relativos à mesma pessoa ou a um grupo de pessoas (na mesma base de dados ou em duas bases de dados diferentes). Caso um adversário possa constatar (por exemplo, por meio de correlação) que dois registros são atribuídos a um mesmo grupo de indivíduos, mas não consegue destacar indivíduos deste grupo, a técnica oferece resistência contra a individualização, mas não contra a capacidade de ligação; e **(iii)** *inference* ou inferência – possibilidade de deduzir, com significativa probabilidade, o valor de um atributo a partir dos valores de um conjunto de outros atributos[196]. Com base nestes parâmetros, far-se-á a análise de algumas das principais e mais populares técnicas de anonimização de dados, especificando suas fragilidades e pontos altos.

Há duas grandes modalidades de técnicas de anonimização: a randomização e a generalização. Dentro de cada uma delas há subcategorias, algumas das quais serão analisadas

[193] OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. **UCLA Law Review**, v. 57, n. 6, p. 1701-1778, 2010. p. 1716.
[194] *Ibidem*. p. 1707.
[195] Uma base de dados é composta por vários registros e cada um deles é correspondente a determinada pessoa (titular dos dados). Um registro é constituído por um conjunto de valores (por exemplo, "Silva") para cada tipo de atributo (por exemplo, "Sobrenome"). Vide: EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 13. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.
[196] *Ibidem*. p. 11-12.

no presente subcapítulo. Outro grupo que merece ser destacado é o das técnicas de supressão, que podem ser totais ou parciais.

3.2.1 Randomização

A randomização é uma família de técnicas de anonimização que também pode ser chamada de "aleatorização" ou "perturbação". A ideia central é modificar os valores de cada atributo para que conste um valor aleatório; altera-se a veracidade dos dados[197]. Quando a exatidão de determinado atributo for crucial para alcaçar a finalidade pretendida com o tratamento, a randomização não deve ser utilizada[198]. Esta modalidade é mais aplicada e eficaz quando o valor é representado por algum número, por exemplo, em datas de nascimento, idade, peso, etc. Outro ponto que deve ser observado, é que o grau de perturbação deve ser proporcional à gama de valores do atributo, isso é, se a base de dados for demasiado pequena, o efeito de anonimização será mais fraco; por outro lado, se a base de dados for demasiado grande, os valores finais serão muito diferentes dos originais e a utilidade do conjunto de dados será provavelmente reduzida[199]. A não observação deste aspecto pode tornar falha a implementação da técnica. Foi o que aconteceu no caso estudado por Narayanan e Shmatikov, envolvendo o concurso promovido pela Netflix. A empresa alegou ter aplicado uma técnica de randomização aos dados, todavia, os pesquisadores constataram que a anonimização foi insuficiente. Os resultados do estudo indicaram que:

> [...] a alegação da Netflix de que os dados estavam perturbados não parece ser confirmada. Um dos assinantes teve 1 de 306 classificações alteradas, e o outro teve 5 de 229 alteradas. [...] Em qualquer caso, o nível de ruído é demasiado pequeno para afetar os algoritmos de desanonimização. Não há como determinar quantas datas foram

[197] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 13. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.
[198] PERSONAL DATA PROTECTION COMMISSION OF SINGAPORE – PDPC. **Guide to basic data anonymisation techniques**. Singapore, 25 jan. 2018. p. 20. Disponível em: <https://iapp.org/resources/article/guide-to-basic-data-anonymization-techniques/>. Acesso em 20 out. 2021.
[199] *Ibidem*.: "The degree of perturbation should be proportionate to the range of values of the attribute. If the base is too small, the anonymisation effect will be weaker; on the other hand, if the base is too large, the end values will be too different from the original and utility of the dataset will likely be reduced*."*.

> alteradas e quantas classificações foram eliminadas, mas suspeitamos que o grau de perturbação aplicada tenha sido deficiente.[200]

Além de atentar para a gama de valores, importante notar que o agir desta modalidade tem como foco o risco de inferência, o qual é consideravelmente diminuído com a sua aplicação, porém, em relação aos outros riscos, a randomização não se mostra muito eficaz. O risco da distinção, por exemplo, não é reduzido por meio da randomização, porquanto os registros continuarão sendo associados a um determinado indivíduo, ainda que os valores não serão fiáveis. O WP29 indica que "pode ser necessária a aplicação de técnicas suplementares para garantir que um registro não seja passível de identificar um indivíduo em particular"[201].

A adição de ruído *(noise addition*) é uma técnica de randomização aplicada a dados numéricos, em que é adicionado um valor aleatório ε (o "ruído"), que pode ser positivo ou negativo, a todos os valores do atributo a ser protegido[202]. Suponha-se que o atributo seja "altura" e que a unidade de medida seja centímetro. Ao invés de constar a altura originalmente medida de um indivíduo, acrescenta-se +-10cm àquela[203]. Quanto maior a variação entre os valores, maior será o nível de randomização. Não é indicado a utilização desta técnica quando os valores originais possuem variações muito significativas. Utilizando-se do exemplo anterior, se um adversário souber que um indivíduo tem uma altura muito maior do que os demais, ele poderá identificar este indivíduo no registro que passou por perturbação e, até mesmo, fazer um palpite plausível da altura original do indivíduo. Por meio deste palpite, ele poderá descobrir a variância e, consequentemente, o valor do "ruído"[204]. Esta é uma das razões pelas quais se recomenda o uso da adição de ruído sempre combinada a

---

[200] "Netflix's claim that the data were perturbed does not appear to be borne out. One of the subscribers had 1 of 306 ratings altered, and the other had 5 of 229 altered. (These are upper bounds, because they include the possibility that the subscribers changed the ratings after the snapshot that was released was taken.) In any case, the level of noise is far too small to affect the deanonymization algorithms. We have no way of determining how many dates were altered and how many ratings were deleted, but we suspect that very little perturbation has been applied.". NARAYANAN, Arvind; SHMATIKOV, Vitaly. How To Break Anonymity of the Netflix Prize Dataset. **ArXiv e-prints**, out. 2006. p. 16 (tradução livre). Disponível em: <https://arxiv.org/abs/cs/0610105#>. Acesso em 25 out. 2021.

[201] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 12. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

[202] UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. p. 96. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021.

[203] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 12. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

[204] UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. p. 96. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021.

alguma outra técnica de anonimização. Um dos erros mais comuns dos controladores é assumir que esta técnica, *per se*, é suficiente. Conforme assinala o WP29, "a adição de ruído é uma medida complementar que torna mais difícil para um adversário recuperar os dados pessoais. [...] não se deve assumir que a adição de ruído representa uma solução autônoma para a anonimização"[205].

Pertence, também, à modalidade da randomização a permutação (*swapping*), que nada mais é do que "embaralhar" os valores de determinado atributo constante em uma base de dados[206]. O valor relativo ao indivíduo X é associado ao indivíduo Y, e assim por diante. Esta técnica é utilizada quando se pretende manter a distribuição exata dos valores (por exemplo, quando é necessária a obtenção da média dos valores para alcançar a finalidade visada com o tratamento dos dados)[207]. A fração correspondente ao número de valores permutados chama-se taxa de permuta (*swap rate*), e é representada por *r*. Geralmente, *r* é da ordem de 1-10%, de modo que a fração de valores trocados será normalmente inferior a um em cada dez[208]. Em alguns casos, são impostas condições sobre quais pares de valores podem ser trocados, o que restringe o número de indivíduos que podem ter valores associados a si permutados. Essas condicionantes normalmente recaem sobre o atributo, motivo pelo qual são chamadas de atributos limitantes (*constraining atributes*). Por exemplo, se a finalidade do tratamento é verificar a taxa de desigualdade salarial entre homens e mulheres, o atributo limitante será o gênero do titular dos dados. Neste caso, para que não seja afetado o objetivo do processamento, apenas poderão ser permutados os valores de pessoas do mesmo gênero[209]. Um erro recorrente quando da implementação desta técnica é selecionar um atributo em que os valores guardam vínculos lógicos com outros atributos da base de dados. Se, a título de exemplo, forem escolhidos para aplicar a permuta os valores referentes ao atributo "salário" e houver, na mesma base de dados, os atributos "profissão" e "ano de nascimento", o ganho para a proteção dos dados pessoais não será significativo. Um adversário consegue,

---

[205] "[…] noise addition is a complementary measure that makes it harder for an attacker to retrieve the personal data. Unless the noise is higher than the information contained in the dataset, it should not be assumed that noise addition represents a standalone solution for anonymisation.". EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 13 (tradução livre). Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

[206] *Ibidem*.

[207] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 14. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

[208] UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. p. 92. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021.

[209] *Ibidem*.

facilmente, deduzir a renda de cada indivíduo, levando-se em conta a sua profissão e a sua faixa etária. Assim como a adição de ruídos, a permutação deve ser aplicada como técnica de anonimização complementar, não sendo suficiente a sua aplicação isolada[210].

3.2.2 Generalização

O pressuposto central da generalização é a modificação da escala ou da ordem de grandeza para que seja feita a substituição dos valores originais por valores menos específicos[211-212]. Para um atributo categórico, várias categorias são combinadas para formar novas categorias menos específicas (por exemplo, ao invés de constar o endereço de uma cidade, constará apenas o país); para um atributo numérico, os valores numéricos são substituídos por intervalos (por exemplo, ao invés de constar em um atributo relativo à idade dos indivíduos o valor "16 anos", constará "menor de 18 anos")[213]. Os valores menos específicos englobam um número maior de pessoas, diminuindo o risco de distinção. Nesse sentido, ainda que os valores originais apresentem significativas variações entre si, a generalização é eficaz, já que, de certa forma, os valores acabam sendo padronizados.

O *k*-anonimato (*k-anonymity*) é uma técnica de generalização por meio da qual cada registro passa a ser indistinguível de pelo menos *k - 1* outros registros[214]. "Diz-se que uma base de dados satisfaz o *k*-anonimato para *k>1* se, para cada combinação de valores de atributos quase-identificadores, existirem pelo menos *k* registros no conjunto de dados compartilhando a mesma combinação"[215]. O valor numérico de *k* indica quantos grupos de registros idênticos serão criados a partir da aplicação da técnica. O adversário fica

---

[210] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 14. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.
[211] SWEENEY, Latanya. Achieving k-anonymity privacy protection using generalization and suppression. International **Journal of Uncertainty, Puzziness and Knowledge-Based Systems**, v. 10, n. 5, p. 557-570, 2002. p. 572.
[212] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 16. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.
[213] PERSONAL DATA PROTECTION COMMISSION OF SINGAPORE – PDPC. **Guide to basic data anonymisation techniques**. Singapore, 25 jan. 2018. p. 62. Disponível em: <https://iapp.org/resources/article/guide-to-basic-data-anonymization-techniques/>. Acesso em 20 out. 2021.
[214] NARAYANAN, Arvind; SHMATIKOV, Vitaly. How To Break Anonymity of the Netflix Prize Dataset. **ArXiv e-prints**, out. 2006. p. 4. Disponível em: <https://arxiv.org/abs/cs/0610105#>. Acesso em 25 out. 2021.
[215] "A data set is said to satisfy k-anonymity for k>1 if, for each combination of quasi-identifier attribute values, at least k records exist in the data set sharing that combination.". EUROPEAN UNION AGENCY FOR NETWORK AND INFORMATION SECURITY - ENISA. **Privacy by design in big data: An overview of privacy enhancing technologies in the era of big data analytics**. Heraklion, dez. 2015. p. 30 (tradução livre). Disponível em: <https://www.enisa.europa.eu/publications/big-data-protection>. Acesso em 20 out. 2021.

incapacitado de distinguir o titular dos dados "em meio à multidão", pois a combinação de quase-identificadores por este apresentada é a mesma que a de vários outros indivíduos representados na base de dados[216]. O objetivo é, por meio da redução a quase zero do risco de distinção do indivíduo, mitigar o risco de possibilidade de ligação. O mesmo registro pode ser relacionado a vários indivíduos mesmo que a base de dados esteja diretamente ligada (ou combinada) a informações externas[217].

Esta técnica foi proposta pela primeira vez em 1998, por Latanya Sweeney e Pierangela Samarati, para garantir a privacidade de titulares que tivessem os seus dados publicizados ou compartilhados com terceiros[218]. As pesquisadoras acreditavam que a retirada de identificadores explícitos (nome, número de celular, endereço,...), técnica geralmente aplicada à época, não era suficiente para proteger os indivíduos[219]. Os titulares de dados ainda poderiam ser distinguidos e identificados por meio da combinação única de quase-identificadores, o que era alcançável pela ligação a informações constantes em outras bases de dados disponíveis publicamente[220]. Levando em conta estes pressupotos, Sweeney e Samarati procuraram desenvolver uma técnica de anonimização que visasse, justamente, impedir estas ligações. Assim, o *k*-anonimato além de providenciar a supressão de parte dos dados, combina a esta a generalização. Para facilitar a visualização do funcionamento desta técnica, traz-se o seguinte exemplo: aos dados do Quadro 4, na qual foram apenas removidos os identificadores explícitos, é aplicado o *k*-anonimato, em que *k=2*, dando origem ao Quadro 5. Os atributos generalizados são "idade" e "CEP", onde os registros são particionados em dois grupos indistinguíveis. O primeiro grupo indistinguível consiste nos registros das linhas 1, 3, 4 e 6 e o segundo é formado pelos registros das linhas 2 e 5[221]:

---

[216] SWEENEY, Latanya. Achieving k-anonymity privacy protection using generalization and suppression. International. **Journal of Uncertainty, Puzziness and Knowledge-Based Systems**, v. 10, n. 5, p. 557-570, 2002. p. 572.
[217] *Ibidem*.
[218] SAMARATI, Pierangela; SWEENEY, Latanya. **Protecting Privacy when Disclosing Information: *k*-Anonymity and Its Enforcement through Generalization and Suppression**. Tech Report, 1998. Disponível em: <http://www.csl.sri.com/papers/sritr-98-04/>. Acesso em 20 out. 2021.
[219] O que foi, inclusive, comprovado pela própria Sweeney no ano de 2000, quando esta realizou o estudo com base nos dados do Censo dos EUA de 1990, já relatado no presente capítulo. Vide: SWEENEY, Latanya. Simple demographics often identify people uniquely. **Health (San Francisco)**, v. 671, n. 2000, p. 1-34, 2000.
[220] SAMARATI, Pierangela; SWEENEY, Latanya. **Protecting Privacy when Disclosing Information: *k*-Anonymity and Its Enforcement through Generalization and Suppression**. Tech Report, 1998. p. 1. Disponível em: <http://www.csl.sri.com/papers/sritr-98-04/>. Acesso em 20 out. 2021.
[221] LIU, Junqiang. Privacy Preserving Data Publishing: Current Status and New Directions. **Information Technology Journal**, v. 11, n. 1, p. 1-8, 2012. p. 2-3.

Quadro 4 – Base de dados pessoais.

IDADE	GÊNERO	CEP	DOENÇA
25	Masc	04848-529	AIDS
25	Fem	04849-333	Pressão Alta
26	Masc	04848-529	Câncer
27	Masc	04894-017	Asma Crônica
27	Fem	04849-333	Pressão Alta
28	Masc	04848-529	Tuberculose

Fonte: Adaptado a partir de LIU, Junqiang. Privacy Preserving Data Publishing: Current Status and New Directions. **Information Technology Journal**, v. 11, n. 1, p. 1-8, 2012. p. 2-3.

Quadro 5 – Base de dados anonimizados por meio de *k*-anonimato.

IDADE	GÊNERO	CEP	DOENÇA
25 - 28	Masc	04848-529 - 04894-017	AIDS
25 -28	Fem	04849-333	Pressão Alta
25 - 28	Masc	04848-529 - 04894-017	Câncer
25 – 28	Masc	04848-529 - 04894-017	Asma Crônica
25 - 28	Fem	04849-333	Pressão Alta
25 - 28	Masc	04848-529 - 04894-017	Tuberculose

Fonte: Adaptado a partir de LIU, Junqiang. Privacy Preserving Data Publishing: Current Status and New Directions. **Information Technology Journal**, v. 11, n. 1, p. 1-8, 2012. p. 2-3.

Apesar de ter bons resultados na mitigação dos riscos de distinção e possibilidade de ligação, o principal defeito do *k*-anonimato é não impedir qualquer tipo de ataques de inferência. Se todos os indivíduos *k* se encontrarem em um mesmo grupo, e se for sabido a que grupo pertence uma determinada pessoa, recuperar o valor referente a esta não se torna uma tarefa difícil. Para reduzir o risco de inferência, o valor numérico de *k* não pode ser pequeno (se *k=2,* então a probabilidade de dois indivíduos compartilharem o mesmo valor é maior do que para *k>10*) [222].

[222] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 17. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

3.2.3 Supressão

A supressão é a eliminação total ou parcial de determinadas informações de uma base de dados. Existem diversos tipos de supressão, alguns deles são: **(i)** supressão de registros – todos os registros relativos a determinado indivíduo são suprimidos da base de dados; **(ii)** supressão de valores – remoção de um determinado valor de atributo em toda a base de dados (quando, por exemplo, é suprimido o nome de todos os titulares); **(iii)** supressão de células – apenas algumas instâncias da tabela são removidas (técnica normalmente utilizada no *k*-anonimato para retirar os valores que "fujam" do padrão estabelecido)[223].

A aplicação exclusiva da supressão, sem combiná-la com outras técnicas de anonimização, apresenta grande falibilidade ao proteger os titulares de dados, porquanto o risco de possbilidade de ligação e de inferência permanecem. Um erro recorrente envolvendo essa modalidade é o fato de muitos acreditarem que a supressão dos identificadores explícitos é suficiente para tornar anonimizada uma base de dados. Conforme incansavelmente demonstrado por Sweeney, esta premissa não procede. Em vários de seus estudos, mediante a ligação de bases de dados (*linking attack*) que tiveram apenas identificadores explícitos suprimidos, a pesquisadora reidentificou os titulares por meio das chamadas "combinações únicas"[224]. A suposta anonimização aplicada ao Censo dos EUA foi uma dentre inúmeras oportunidades em que Sweeney comprovou a deficiência da aplicação exclusiva da supressão. Sua primeira contribuição foi em 1997, quando, após a divulgação pela agência governamental *Group Insurance Commission* – GIC de registros que compilavam as visitas hospitalares de 135 mil funcionários do estado de Massachusetts, logrou reidentificar os dados referentes ao governador à época, William Weld[225]. Após constatar que Weld residia na cidade de Cambridge, Sweeney comprou, por somente vinte dólares, um boletim com dados eleitorais relativos aos 54 mil cidadãos de tal município. Apesar de o governo estadual ter declarado que os dados divulgados pelo GIC teriam passado por processo de anonimização, posto que não constantes informações como nome e endereço dos pacientes, a pesquisadora demonstrou ser plenamente possível a reidentificação daqueles pela combinação com outras bases de dados (no caso, as informações externas utilizadas na reversão foram os dados

[223] GRACE, Paul et al. **D4.3 – Guidelines for data anonymization report**. Out. 2016. p. 08. Disponível em: <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ade9ca8a&appId=PPGMS>. Acesso em 05 nov. 2021.
[224] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 34. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.
[225] SWEENEY, Latanya. *k*-Anonymity: a model for protecting privacy. **International Journal on Uncertainty, Fuzziness and Knowledge-based Systems**, v. 10, n. 5, p. 557-570, 2002. p. 558.

constantes no boletim de dados eleitorais). Por meio do cruzamento das bases de dados, Sweeney constatou que seis pacientes tinham a mesma data de nascimento que o governador; três deles eram homens; e apenas um (que era o próprio Weld) tinha o mesmo *ZIP-Code*[226].

Outro fator problemático relacionado ao método de supressão é que o *trade-off* utilidade *vs.* anonimização pode ser intensificado com a sua implementação. Quanto mais atributos são retirados de uma base de dados, mais se perde a utilidade desta. Há dois fatores, portanto, que devem ser observados quando da aplicação da supressão: o primeiro é que o seu uso exclusivo pode ser insuficiente para proteger os titulares, uma vez que a combinação com informações externas pode permitir a reidentificação; a segunda é o fato de que, para alcançar um nível satisfatório de proteção, a supressão, *per se*, pode causar a perda de utilidade da base de dados. Por conta destes dois cenários, a recomendação é que a supressão deve ser, preferencialmente, conjugada a outras modalidades de anonimização, o que é proposto por Sweeney e outros pesquisadores que detectaram falhas no uso exclusivo da referida modalidade[227]. Com o fim de facilitar a visualização de como funciona o processo de supressão em conjunto com outra técnica, no caso, a generalização, utilizar-se-á o banco de dados relacional estruturado em tabela que fora apresentado no capítulo anterior (Quadro 1).

Quadro 6 – Banco de dados relacional anonimizado.

NOME	CPF	CEP	IDADE	SEGMENTAÇÃO
Carlos [redacted]	[redacted]	91978-[redacted]	< 18	Jovem jogador de tênis
Carlos [redacted]	[redacted]	42900-[redacted]	< 18	Jovem sedentário
Carlos [redacted]	[redacted]	91978-[redacted]	> 18	Jovem fisiculturista
Carlos [redacted]	[redacted]	42368-[redacted]	> 18	Adulto ciclista

Fonte: Adaptado de BIONI, Bruno. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, jul. 2015. p. 26.

No exemplo acima, foi aplicada a técnica da generalização nos atributos nome, CEP e idade. Em não constando o sobrenome dos indivíduos na base de dados, estes ficam

[226] *Ibidem.* p. 559.
[227] Vide: SWEENEY, Latanya. Achieving k-anonymity privacy protection using generalization and suppression. **International Journal of Uncertainty, Puzziness and Knowledge-Based Systems**, v. 10, n. 5, p. 557-570, 2002. p. 571-588; e KNAPP, Gerry; OROOJI, Marmar. Improving Suppression to Reduce Disclosure Risk and Enhance Data Utility. In: IISE Annual Conference, mai. 2018, Orlando/FL. **Proceedings of the 2018** [...]. Orlando: Institute of Industrial & Systems Engineers (IISE), 2018. p. 1415-1420. Disponível em: <https://arxiv.org/pdf/1901.00716.pdf>. Acesso em 21 set. 2021.

indistinguíveis, porquanto possuem o mesmo nome. Da mesma forma, se forem suprimidos os três últimos dígitos do CEP (supressão parcial de valores), a localização ficaria menos detalhada, quebrando, em tese, o vínculo de identificação desta informação com um sujeito determinado. Para que não fosse utilizada a idade exata, o fornecimento de tal informação passaria por processo de generalização para que constassse apenas a faixa etária dos indivíduos, como maiores ou menores de 18 anos. Em relação ao CPF, por se tratar de identificador único, permitindo a distinção dos sujeitos (inclusive de homônimos) de maneira inequívoca e imediata, a generalização não seria suficiente, razão pela qual, neste atributo, foi aplicada a supressão total dos valores[228].

[228] BIONI, Bruno. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, jul. 2015. p. 25-26.

4. O CRITÉRIO DA RAZOABILIDADE: UM CONCEITO JURÍDICO INDETERMINADO

O crescente número de dados produzidos e disponibilizados pelas mais diversas fontes somado à contínua sofisticação dos algoritmos de análise de dados torna cada vez mais fácil a combinação de bases de dados e a inferência de dados pessoais a partir de dados considerados anonimizados[229]. A reversibilidade do processo de anonimização de dados é uma premissa quase que irrefutável no âmbito da ciência da computação[230]. Tornar-se-ia insustentável ao Direito continuar se baseando na antiga ideia da anonimização robusta, além do que, a proteção garantida ao titular dos dados estaria propensa a grandes falhas. Diante desta realidade, as leis de proteção de dados passaram a adotar abordagens baseadas na ideia de que "a anonimização absoluta de um dado pode ser difícil ou mesmo impossível em certos casos, assim a adequação deve ser avaliada de forma contextual e por meio do grau de risco existente"[231]. As interpretações jurídicas acerca da avaliação do nível de adequação de um processo de anonimização dividem-se em duas correntes: a LGPD e o GDPR adotam a chamada *risk-based approach* ou abordagem baseada no risco, segundo a qual se entende que há um risco inerente de o dado anonimizado se transmudar em dado pessoal, ainda que tenham sido tomadas precauções para evitá-lo[232]; há também a *procedure-based approach* ou abordagem baseada nos procedimentos adotados, em que o foco da análise está na implementação ou não de procedimentos apropriados aos riscos previamente detectados[233].

Assume-se, portanto, que a anonimização não torna impossível a possibilidade de identificação do sujeito ao qual os dados se referem, mas, sim, remota[234]. Dessa forma, para diferenciar o conceito de dado pessal do conceito de dado anonimizado, já que em ambos o titular dos dados seria identificável, as leis que adotam uma abordagem baseada no risco introduziram o critério da razoabilidade. A mera possibilidade de que um dado anonimizado

---

[229] FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020. p. 11.
[230] Vide capítulo 2 da presente monografia.
[231] JÚNIOR, Odélio Porto. **Anonimização e Pseudonimização: conceitos e diferenças na LGPD**. 25 mai. 2019. Disponível em: <https://baptistaluz.com.br/espacostartup/anonimizacao-e-pseudonimizacao-conceitos-e-diferencas-na-lgpd/#_ftn4>. Acesso em 04 out. 2021.
[232] FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020. p. 14.
[233] JÚNIOR, Odélio Porto. **Anonimização e Pseudonimização: conceitos e diferenças na LGPD**. 25 mai. 2019. Disponível em: <https://baptistaluz.com.br/espacostartup/anonimizacao-e-pseudonimizacao-conceitos-e-diferencas-na-lgpd/#_ftn4>. Acesso em 04 out. 2021.
[234] UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. p. 6. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021.

seja atrelado a uma pessoa deixa de ser suficiente para atrair o termo "identificável". A vinculação precisa se dar mediante esforços razoáveis[235]. A razoabilidade é, portanto, um instrumento que permite determinar quando um risco é ou não significativo para fins de reidentificação. A adoção deste parâmetro permitiu que a dicotomia mutuamente excludente entre dados pessoais e dados anonimizados deixasse de ser algo incoerente, uma vez que a redundância normativa antes verificada deixou de existir[236].

A ideia relativa à "utilização de meios razoáveis" passou a vigorar a partir da Diretiva 95/46/EC, e foi desenvolvida com a finalidade de tornar mais concreta a análise da identificabilidade dos dados que passaram por processo de anonimização, evitando que fosse algo puramente hipotético[237]. Assim como se dá no atual Regulamento Europeu, o legislador brasileiro reconhece, expressamente, a possibilidade de reversão do processo de anonimização mediante esforços razoáveis, caso em que as disposições da Lei se aplicam. Se os esforços extrapolarem a razoabilidade, o dado é considerado anonimizado, ficando, pois, fora do escopo de aplicação da LGPD. O artigo 12, *caput*, da Lei assim prevê:

> Art. 12. Os dados anonimizados não serão considerados dados pessoais para os fins desta Lei, salvo quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido.[238]

Para não se manter indiferente aos avanços tecnológicos que podem tornar técnicas de anonimização, que hoje são consideradas eficazes, obsoletas no futuro, as leis de proteção de dados optaram por adotar um conceito jurídico indeterminado. Isso quer dizer que a razoabilidade consiste em um conceito "cujo conteúdo e extensão são em larga medida

---

[235] BIONI, Bruno Ricardo. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 41.

[236] BIONI, Bruno Ricardo. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 42.

[237] DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. *E-book*. p. RB-8.2.

[238] BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

incertos, ou seja, não são dotados de um sentido preciso e objetivo"[239]. Judith Martins-Costa destaca que os conceitos jurídicos indeterminados podem ser divididos em duas espécies: aqueles relativos a realidades valorativas e aqueles relativos a realidades fáticas[240]. A razoabilidade pertence ao primeiro grupo, porquanto integra a descrição do fato em exame (o risco de reversão da anonimização dos dados), ao qual são atribuídas determinadas consequências jurídicas (estar ou não dentro do escopo material de aplicação da Lei)[241].

Conforme assevera José Alfredo de Oliveira Baracho, o problemático de aplicar um conceito indeterminado, com grau significativo de abstração, é que isto gera uma pluralidade de opiniões sustentáveis mediante argumentos lógicos, que se desdobram da pura interpretação jurídica, para se estender a juízos de tipo técnico ou de puras valorações fáticas[242]. Assim, com a introdução do critério da razoabilidade, ao mesmo tempo que o legislador promove a neutralidade da Lei em relação à tecnologia[243], impedindo que aquela se torne defasada com o tempo, abre-se espaço para a discricionariedade quando da interpretação da norma. Por conta da abertura semântica do conceito, a determinação do que é ou não razoável pode provocar certa insegurança àqueles que aplicam, bem como aos que regulam a implementação da anonimização em bases de dados pessoais. Apesar de não estabelecer parâmetros específicos (como valores numéricos), a LGPD prescreveu balizas para reduzir a discricionariedade do exercício interpretativo e, com isso, alcançar um mínimo de previsibilidade à aplicação da norma[244]. Estes critérios encontram-se previstos no § 1º do artigo 12 da Lei:

> § 1º A determinação do que seja razoável deve levar em consideração fatores objetivos, tais como custo e tempo necessários para reverter o processo de anonimização, de acordo com as tecnologias disponíveis, e a utilização exclusiva de meios próprios.[245]

---

[239] ROZAS, Luiza Barros. Conceitos jurídicos indeterminados e discricionariedade administrativa. **Cadernos Jurídicos da Escola Paulista de Magistratura**, v. 20, n. 47, p. 191-201, 2019. p. 192.
[240] MARTINS-COSTA, Judith. **A boa-fé no direito privado**: critérios para a sua aplicação. 2. ed. São Paulo: Saraiva Educação, 2018. *E-book*. p. 156-157.
[241] MARTINS-COSTA, Judith. **A boa-fé no direito privado**: critérios para a sua aplicação. 2. ed. São Paulo: Saraiva Educação, 2018. *E-book*. p. 158.
[242] BARACHO, José Alfredo de Oliveira. Teoria geral dos conceitos legais indeterminados. **Themis**, v. 2, n. 2, p. 61-78, 1999. p. 61-62.
[243] BIONI, Bruno Ricardo. **Calibrando o filtro da razoabilidade: critérios objetivos e subjetivos como fatores de uma análise de risco**. 14 nov. 2019. Disponível em: <http://genjuridico.com.br/2019/11/14/filtro-da-razoabilidade-criterios/>. Acesso em 27 out. 2021.
[244] *Ibidem.*
[245] BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

Consoante se depreende do citado dispositivo, a delimitação do que se entende por "razoável" se dá por meio da verificação de dois eixos distintos: o eixo subjetivo (utilização exclusiva de meios próprios para a reversão) e o eixo objetivo (custo econômico e tempo exigidos para reverter o processo de anonimização, consoante o estado da arte da tecnologia). Se os dados não passarem no "teste da razoabilidade", eles estarão fora do escopo de aplicação do regime de proteção de dados pessoais, sendo considerados anonimizados[246]. Tanto no aspecto objetivo quanto no subjetivo, demanda-se a realização de uma análise contextual, o que faz do "status" de dado pessoal um estado dinâmico que pode ser alterado a depender das circunstâncias encontradas caso a caso[247].

Antes da aprovação da LGPD, Bruno Bioni realizou, no ano de 2016, uma análise das iniciativas legislativas que visavam regular a proteção de dados pessoais no país[248]. Na oportunidade, Bioni apontou três diferentes estratégias regulatórias que poderiam ser adotadas para obter a mensuração do que viesse a ser um risco razoável/aceitável de reidentificação: **(i)** silêncio; **(ii)** regulação *ex ante*; **(iii)** regulação *ex post*. Se a Lei apenas enunciasse a razoabilidade, permanecendo-se silente em relação aos critérios de interpretação, os atores se autorregulariam. Com a regulação *ex ante*, o legislador indicaria os critérios a serem considerados para aferir o que viesse a ser razoável. Neste cenário, poderia haver uma autorregulação parcial ou uma heterorregulação parcial, isso é, os atores não seriam absolutamente livres para se autorregular, uma vez que o órgão regulador atuaria para reduzir a discriscionariedade na interpretação da norma. Por fim, a regulação *ex post* promoveria uma heterorregulação total, pois os critérios seriam indicados, exclusivamente, pelo órgão regulador[249]. Com a Lei já em vigor, verifica-se que o legislador optou pela regulação *ex ante*, em que se mescla a autorregulação dos atores ao papel fiscalizador da Autoridade Nacional de Proteção de Dados - ANPD. Ainda que a LGPD discipline os critérios a serem levados em conta no "teste da razoabilidade", o § 3º do artigo 12 estabelece que quem tem a competência para dispor sobre padrões e técnicas de anonimização é a ANPD:

---

[246] FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020. p. 13.
[247] PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. **Law, Innovation and Technology**, v. 10, n. 1, p. 41-81, 2018. p. 47: "The resulting standard of the reasonable likelihood of identification is quite broad and context-dependent, leading to one major consequence: the status of data as 'personal' is dynamic (…)".
[248] BIONI, Bruno Ricardo. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, 2015. p. 34-35.
[249] *Ibidem.*

> § 3º A autoridade nacional poderá dispor sobre padrões e técnicas utilizados em processos de anonimização e realizar verificações acerca de sua segurança, ouvido o Conselho Nacional de Proteção de Dados Pessoais.[250]

Conforme se depreende do texto da Lei, cumpre à ANPD dispor sobre os procedimentos padrões a serem adotados para fins de anonimização. Todavia, sabe-se que, na prática, aqueles agentes de tratamento que visarem aplicar técnicas de anonimização de dados deverão, eles próprios, avaliar se os instrumentos adotados estão de acordo com os padrões estabelecidos e se são suficientes para, de fato, anonimizar os dados e proteger os titulares. A ANPD não fornece qualquer assistência prática em termos de ajudar as organizações a determinar se os dados por elas anonimizados são suscetíveis de resultar em reidentificação de titulares (seria, inclusive, inviável)[251]. O mesmo acontece em relação ao GDPR e, levando-se isso em consideração, o *Information Commissioner's Officer* – ICO (órgão que possui papel análogo à ANPD) do Reino Unido, sugere, em parecer por ele elaborado, que as organizações contratem profissionais que realizem testes para verificar a adequação dos processos de anonimização[252]. O chamado "*the motivated intruder test*", no português, teste do adversário motivado, baseia-se na ideia de que um adversário, que começa sem nenhum conhecimento prévio, deseja identificar o titular dos dados que passaram pelo processo de anonimização a ser testado. Assume-se que o adversário tenha acesso a meios razoáveis, como *internet*, bibliotecas e registros públicos, e que poderia adotar técnicas investigativas, como falar com pessoas que poderiam fornecer informações adicionais[253]. O teste, portanto, destina-se a avaliar se o adversário seria bem sucedido para, então, atestar a qualidade da anonimização aplicada pela organização. Avaliações como esta relatada poderiam ser, da mesma forma, recomendadas pela ANPD, o que proporcionaria maior segurança aos agentes de tratamento quando da aplicação de técnicas de anonimização.

---

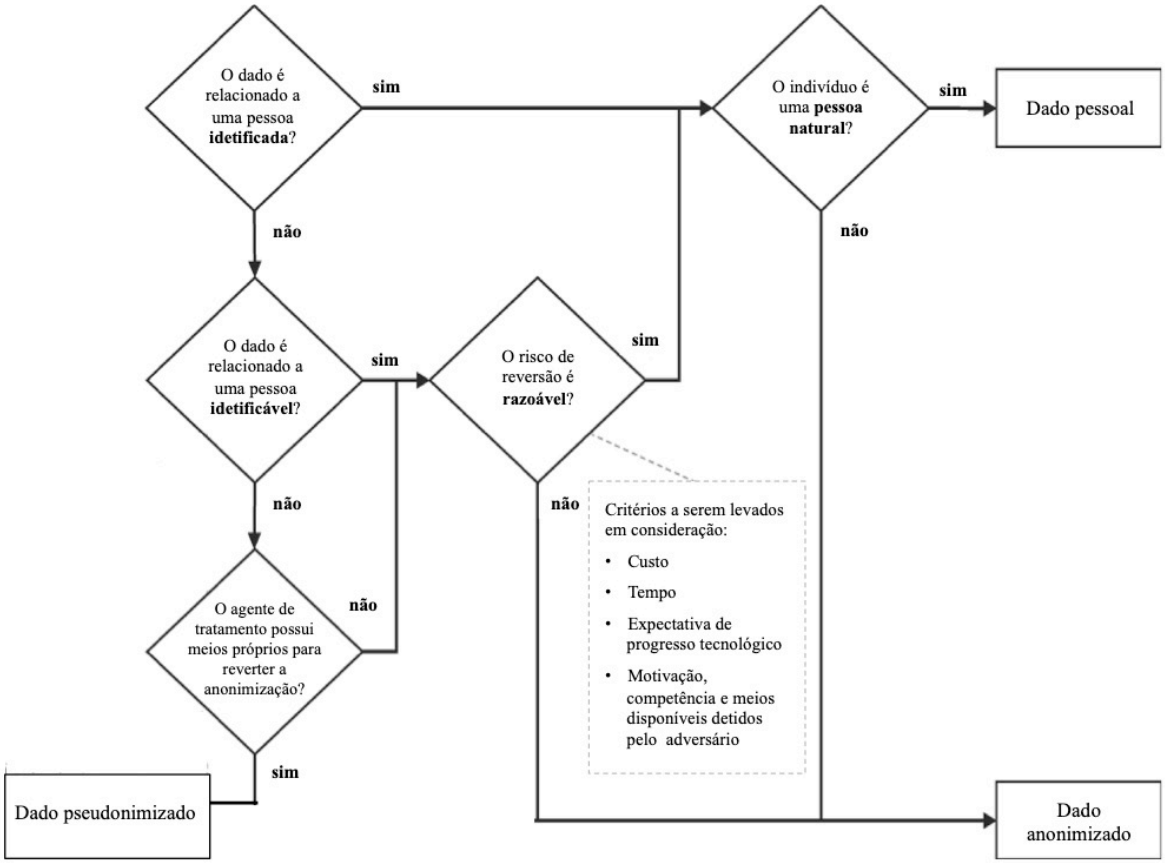
[250] BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.
[251] UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. p. 22. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021.
[252] *Ibidem.*
[253] *Ibidem,* p. 22-23.

4.1 O TESTE DA RAZOABILIDADE

A aplicação do critério da razoabilidade para distinguir dados pessoais de dados anonimizados depende de uma análise a ser realizada caso a caso. O contexto e as circunstâncias da situação específica desempenham papel fundamental em tal avaliação[254]. Assim, aplica-se um teste geral a fim de analisar todos os fatores em causa. A representação esquemática abaixo indica as fases do teste da razoabilidade, as quais se dividem com base nos parâmetros estabelecidos na LGPD. O primeiro ponto a se verificar é se o dado é referente a uma pessoa identificada ou a uma pessoa identificável. Se a pessoa à qual o dado se refere for identificada e, desde que seja uma pessoa natural, o dado será categorizado como pessoal. Por sua vez, se o dado for relativo à pessoa identificável, e tiver passado por processo de anonimização, passar-se-á à avaliação dos eixos subjetivo e objetivo. No eixo subjetivo, verifica-se se o agente de tratamento possui meios próprios para reverter o processo de anonimização. Se a resposta for positiva, os dados serão pseudonimizados e, dessa forma, são considerados pessoais e estarão dentro do escopo de aplicação da LGPD. Em sendo a resposta negativa, serão avaliados os critérios do eixo objetivo. Além do custo, do tempo e do estado da arte das tecnologias disponíveis, que estão previstos no corpo da Lei, devem ser analisados o volume e a natureza dos dados, assim como as motivações, a competência e os meios disponíveis que podem ser detidos pelo adversário. Neste último, a análise pode se dar por meio do antes referido teste do adversário motivado, em que um terceiro, contratado pela organização, testa o processo de anonimização implementado, agindo como se o adversário fosse para tentar reverter a anonimização.

[254] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques**. Brussels, 10 abr. 2014. p. 8. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

Figura 3 – Teste da razoabilidade.



Fonte: Adaptado a partir de FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020. p. 14.

Passa-se, agora, a uma abordagem específica de ambos os planos em que a análise do critério da razoabilidade se divide: o eixo subjetivo e o eixo objetivo.

4.1.1 Eixo subjetivo

No eixo subjetivo, a análise recai sobre os meios próprios do controlador ou do operador, isso é, dos agentes de tratamento dos dados[255]; foca-se na capacidade de um agente em específico[256]. Deve-se avaliar, caso a caso, se com os recursos detidos por aquele que processa os dados, este é capaz de reverter o processo de anonimização. Em sendo possível a

[255] DATAGUIDANCE BY ONETRUST; BAPTISTA LUZ ADVOGADOS**. Comparing privacy laws: GDPR v. LGPD**. 2019. Disponível em: <https://baptistaluz.com.br/wp-content/uploads/2019/05/DataGuidance-GPDR-LGPD-For-Print.pdf>. Acesso em 28 out. 2021.

[256] BIONI, Bruno Ricardo. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 47.

reversão, estar-se-á diante de dados pseudonimizados, os quais se encontram dentro do escopo de aplicação da LGPD, visto que são considerados dados pessoais.

É comum que as organizações se divididam em segmentos a depender do tipo de atividade desenvolvida (por exemplo, em um escritório de advocacia, há o setor administrativo, o setor de informática, o setor de *marketing*...). Ainda que determinado setor tenha somente acesso a uma base de dados anonimizada, o fato de outro setor da mesma empresa deter a base de dados pessoais original já é suficiente para alterar a qualificação dos dados da base supostamente anonimizada para dados pseudonimizados[257]. O agente de tratamento detém informações adicionais para, ele mesmo, reverter o processo de anonimização. Todavia, se a base de dados original for eliminada, o processador dos dados não mais poderá realizar a reversão por meios próprios. A base de dados será, em tese, anonimizada, porém, necessária a análise do ponto de vista objetivo para que se tenha certeza.

Um caso conhecido que envolve a manutenção de meios próprios pelo agente de tratamento e que, mesmo assim, foi erroneamente considerado por este como anonimização é o dos Táxis da cidade de Nova York. Anualmente, a Comissão de Táxi e Limousine (*Taxi and Limousine Commission* - TLC) da cidade divulga bancos de dados em que constam atributos como horário, coordenadas geográficas, tarifa e valor de gorjeta de aproximadamente 173 milhões de viagens individuais. Após a publicização destes dados, no ano de 2013, foram descobertas as falhas do sistema de proteção implementado e, por meio da combinação com informações provenientes de outras fontes (como fotos de *paparazzi* divulgadas na Internet), foi possível, por exemplo, obter as rotas de viagens de diversas celebridades que fizeram o uso do serviço, tais como Bradley Cooper e Olivia Munn[258]. Além dos dados relativos aos passageiros, aqueles referentes aos motoristas dos táxis também foram disponibilizados. Para que se mantesse oculto o valor do salário de cada profissional, aplicou-se anonimização ao número de licença do motorista e ao número do táxi[259]. Todavia, apesar de estes identificadores explícitos terem sido anonimizados, os quase-identificadores não o foram, o que permitiu a reidentificação dos motoristas por meio da combinação com outras bases de

---

[257] BIONI, Bruno Ricardo. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 44.

[258] PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. **Law, Innovation and Technology**, v. 10, n. 1, p. 41-81, 2018. p. 47.

[259] DOURIEZ, Marie (et. al.). Anonymizing NYC Taxi Data: Does It Matter?. In: IEEE International Conference on Data Science and Advanced Analytics (DSAA), out. 2016, Montréal/QC. **Proceedings of 2016 (…)**. Montréal: IEEE, 2016. p. 140-148. Disponível em: <https://ieeexplore.ieee.org/document/7796899>. Acesso em 27 out. 2021.

dados. O risco de possibilidade de ligação foi mantido, mesmo após a implementação da técnica de anonimização. Outra falha constatada, foi o fato de que a codificação alfanumérica atribuída aos dois identificadores explícitos foram mantidas em todos os registros, ou seja, o mesmo táxi e o mesmo motorista possuíam, cada um, sempre o mesmo código, o que possibilitou, assim a sua distinção[260].

Na Europa, desenvolve-se uma discussão jurisprudencial e doutrinária a respeito da (ir)relevância dos conhecimentos e dos meios detidos por terceiros para reverter o processo de anonimização. Os posicionamentos confrontantes dividem-se em duas teorias: a teoria relativa e a teoria absoluta[261 - 262]. Segundo a primeira delas, o teste da razoabilidade se dá apenas em relação aos meios e conhecimentos detidos pelo agente de tratamento dos dados. Já na teoria absoluta, além de se considerar o plano do responsável pelo processamento das informações, leva-se em conta, também, os meios e conhecimentos detidos por terceiros[263]. Ainda que em alguns países, como na Alemanha, a teoria relativa seja preferida, o Tribunal de Justiça da União Europeia – TJUE reconheceu, no paradigmático acórdão *Breyer*, a relevância jurídica dos conhecimentos e meios detidos por terceiros, o que fica claro na seguinte passagem:

> [...] um endereço de protocolo Internet dinâmico registrado por um prestador de serviços de meios de comunicação em linha quando da consulta por uma pessoa de um sítio Internet que esse prestador disponibiliza ao público constitui, relativamente a esse prestador, um dado pessoal na ação dessa disposição, quando este disponha de meios legais que lhe permitam identificar a pessoa em causa graças às informações suplementares que o fornecedor de acesso à Internet dessa pessoa dispõe.[264]

---

[260] PANDURANGAN, Vijay. **On taxis and rainbows: Lessons from NYC's improperly anonymized taxi logs**. Medium. 21 jun. 2014. Disponível em: <https://tech.vijayp.ca/of-taxis-and-rainbows-f6bc289679a1>. Acesso em 27 out. 2021.

[261] A teoria absoluta pode ser, também, denominada de teoria objetiva. Vide: CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020.

[262] SPINDLER, Gerald; SCHMECHEL, Philipp. Personal Data and Encryption in the European General Data Protection Regulation. **Journal of Intellectual Property, Information Technology and Electronic Commerce Law**, v. 7, n. 2, p. 163-177, 2016. p. 165-166.

[263] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 126.

[264] UNIÃO EUROPEIA. TJUE. Processo n. C-582/14 - Breyer. Autor: Patrick Breyer. Réu: Bundesrepublik Deutschland. 19 out. 2016. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A62014CJ0582>. Acesso em 27 out. 2021.

No Considerando 26 do GDPR, consta, de forma expressa, que a análise da identificabildidade deve levar em conta todos os meios razoavelmente prováveis de serem utilizados, tanto pelo controlador quanto por outra pessoa. Fica evidente que o legislador europeu optou pela teoria absoluta, já que leva em conta os meios detidos por teceiros ("ou por outra pessoa"). Em relação à LGPD, pode-se presumir que, em razão do uso da conjunção alternativa "ou" no *caput* do artigo 12, o legislador brasileiro é, da mesma forma, adepto à teoria absoluta. Todavia, não fica claro se o termo "meios próprios" considera ou não o fluxo de dados para fora da organização responsável pelo tratamento. No caso, por exemplo, de determinada empresa possuir parcerias que envolvam o uso compartilhado de dados[265 - 266], fica vago se a análise subjetiva estender-se-ia ao âmbito da empresa parceira (a qual poderia ser considerada análoga a um setor da outra empresa, por exemplo) ou se recairia apenas sobre os meios próprios detidos pela empresa "principal".

O supracitado acórdão *Breyer* tratou de outra temática polêmica envolvendo o eixo subjetivo do teste da razoabilidade: se são ou não considerados como meios para a reidentificação aqueles obtidos de forma ilícita. Segundo o entendimento do TJUE ratificado na referida decisão, "somente fatores lícitos devem ser considerados para efeitos de preenchimento do elemento da identificabildiade"[267]. De acordo com o Tribunal Europeu, as informações suplementares que venham a possibilitar a reidentificação do titular dos dados não poderão ser consideradas se forem obtidas ilicitamente[268]. Este posicionamento é criticado por parte da doutrina europeia. Os defensores da possibilidade de os meios serem ilícitos assumem que "sendo o critério último o da razoabilidade, importa atender a todas as

---

[265] Art. 5º [...] XVI - uso compartilhado de dados: comunicação, difusão, transferência internacional, interconexão de dados pessoais ou tratamento compartilhado de bancos de dados pessoais por órgãos e entidades públicos no cumprimento de suas competências legais, ou entre esses e entes privados, reciprocamente, com autorização específica, para uma ou mais modalidades de tratamento permitidas por esses entes públicos, ou entre entes privados; [...]. BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

[266] BIONI, Bruno Ricardo. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 46.

[267] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 123.

[268] UNIÃO EUROPEIA. TJUE. Processo n. C-582/14 - Breyer. Autor: Patrick Breyer. Réu: Bundesrepublik Deutschland. 19 out. 2016. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A62014CJ0582>. Acesso em 27 out. 2021: "Como salientou o advogado geral, em substância, no n.o 68 das suas conclusões, assim não será se a identificação da pessoa em causa for proibida por lei ou inexequível, por exemplo devido ao facto de implicar um esforço desmedido em termos de tempo, de custo e de mão de obra, de modo que o risco de uma identificação parece na realidade insignificante.".

condutas ilícitas desde que se possa razoavelmente contar com elas"[269]. O legislador brasileiro não se posiciona a respeito deste aspecto, o qual tampouco foi objeto de análise do Judiciário. A lacuna pode ser, como já realizado em outras oportunidades, preenchida consoante a experiência europeia.

4.1.2 Eixo objetivo

A solidez de determinada técnica de anonimização é verificada por meio de três riscos que podem propiciar a reidentificação do titular: o risco de distinção, a possibilidade de ligação e o risco de inferência. Como pôde se observar na análise realizada no capítulo anterior, o risco de distinção ou *singling out* constitui o risco de mais fácil eliminação. A possibilidade de ligação e o risco de inferência, contudo, demandam maior grau de atenção. De acordo com o *Information Commissioner's Office* - ICO do Reino Unido, o risco de reidentificação do titular por meio da combinação de bases de dados é imprevisível, pois não há como apurar quais dados estão disponíveis publicamente ou quais serão disponibilizados no futuro[270]. Por conta da remanescência, essencialmente, dos riscos de possibilidade de ligação e de inferência, adotam-se critérios objetivos a serem observados que afastam a categorização do dado como "pessoal" se a reidentificação, seja por meio de tecnologias que permitam a ligação a outras bases de dados ou a inferência de valores, se dê de maneira muito morosa ou custosa.

Pode-se dizer que o eixo objetivo de análise atém-se a padrões sociais, porquanto demanda a verificação de como o estado da técnica calibra a escala de recursos (tempo e custo) para transmudar um dado anonimizado em dado pessoal[271]. Trata-se de fator de capacidade geral, aplicável indistintamente a todos, e não de forma individualizada. Porém, em situações especiais, o tempo e o custo econômico que seriam geralmente considerados desarrazoados, podem passar a ser razoáveis. Se, por exemplo, pela notoriedade do titular, os benefícios advindos da reversão da anonimização forem consideráveis para um adversário, o

---

[269] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 123.

[270] UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. p. 18. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021: "[...] the risk of re-identification through data linkage is essentially unpredictable because it can never be assessed with certainty what data is already available or what data may be released in the future.".

[271] BIONI, Bruno Ricardo. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 43.

dispêndio de quantias monetárias e de tempo que de outro modo não seriam razoáveis a ele, passam a ser[272].

A LGPD determina que os meios técnicos razoáveis devem estar disponíveis na ocasião (Art. 5º, III) e no momento (Art. 5º, XI) do tratamento. Delimita-se o estado da arte da tecnologia ao período em que o processamento dos dados é realizado (isso é, enquanto o tratamento perdurar). Isso reforça o fato de que a anonimização não deve ser "aplicada e esquecida". Deve haver um controle contínuo do manejo dos riscos durante todo o "ciclo de vida" do dado[273]. As técnicas implementadas devem ser periodicamente testadas para avaliar se continuam eficazes ou se precisam ser atualizadas para garantir que a identificabilidade dos titulares permaneça sendo remota. A razão disso é que a identificação pode não ser possível com o conjunto dos meios suscetíveis de serem razoavelmente utilizados hoje, mas, a título exemplificativo, se o armazenamento perdurar por 10 anos, o responsável pelo tratamento deverá considerar a possibilidade de identificação que pode surgir no futuro[274]. Conforme assevera o WP29, "o sistema deverá ter a capacidade de se adaptar a estas evoluções à medida que elas ocorrem, permitindo incorporar, em tempo útil, as medidas técnicas e organizacionais que se imponham"[275].

4.2 A ANONIMIZAÇÃO COMO MEDIDA DE MITIGAÇÃO DE RISCOS

Perante o cenário de aprimoramento constante da tecnologia, o que impõe aos agentes de tratamento o cumprimento de obrigações de ordem técnica e de gestão de forma contínua, questiona-se se se é válido a uma organização assumir todo o esforço mencionado sendo que a anonimização sequer torna inviável a possibilidade de reidentificação dos titulares[276]. A resposta é: sim.

---

[272] CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020. p. 125-126.
[273] FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020. p. 16.
[274] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 16. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.
[275] EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. p. 16. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.
[276] BIONI, Bruno Ricardo. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 44.

Em razão de o regime de proteção de dados não se aplicar aos dados anonimizados, vários agentes de tratamento veem na anonimização uma forma de se esquivar das disposições da Lei. Contudo, esse não deve ser o objetivo principal para a implementação de técnicas de anonimização. A principal motivação para tal deveria ser (e para muitos é), essencialmente, a mitigação dos riscos advindos da atividade de tratamento de dados, quer para os titulares de dados, quer para o próprio responsável pelo processamento. Ainda que um conjunto de dados anonimizado guarde sempre um risco residual de reidentificar os titulares das informações, o risco de danos em caso de eventual vazamento ou acesso indevido à base de dados é consideravelmente reduzido se comparado a uma base de dados em "estado bruto". É exatamente por conta disso que a LGPD ao dispor sobre o tratamento de dados pessoais sensíveis, que, como o próprio nome indica, são os dados que mais tornam os titulares vulneráveis, sempre recomenda a implementação de técnicas de anonimização a estas informações[277].

Na era da sociedade da informação são poucas as organizações que não realizam o tratamento de dados pessoais em algum grau para o desenvolvimento de determinadas atividades. Farmácias, universidades, lojas virtuais, escritórios de advocacia, redes sociais... São incontáveis as esferas que desenvolvem uma relação de verdadeira dependência aos dados. O processamento responsável destas informações traz benefícios à sociedade como um todo. Além de impulsionar o desenvolvimento de conhecimento e o avanço da transformação digital, o tratamento de dados pessoais passa a conquistar maior confiança dos titulares de dados quando realizado em consonância com a Lei[278]. No âmbito das instituições privadas, a *compliance* de dados configura um fator de vantagem competitiva no mercado. Com recorrentes vazamentos de dados pessoais noticiados pela imprensa[279], aquelas empresas que adotarem ferramentes de segurança da informação e demonstrarem ao seu público o seu

[277] Vide subcapítulo 2.1 da presente monografia.
[278] ARBIX, Daniel. A importância da privacidade por *design* e por *default* (*privacy by design and by default*). In.: DONEDA, Danilo et. al. (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 55.
[279] Por exemplo: HOFFMANN, Bruno. **Falha da Enel expõe dados de milhões de clientes**. 04 fev. 2020. Disponível em: <https://www.gazetasp.com.br/estado/2020/02/1061629-falha-da-enel-expoe-dados-de-milhoes-de-clientes.html>. Acesso em 27 out. 2021.; LEOCÁDIO, Thais; PIMENTA, Guilherme. **Dados pessoais e jornada de trabalho dos professores da rede estadual vazam na internet**. 14 mai. 2021. Disponível em: <https://g1.globo.com/mg/minas-gerais/noticia/2021/05/14/dados-pessoais-e-jornada-de-trabalho-dos-professores-da-rede-estadual-vazam-na-internet.ghtml>. Acesso em 27 out. 2021.; PAYÃO, Felipe. **Detran SP tem vazamento de 2 milhões de dados**. 16 ag. 2021. Disponível em: <https://www-tecmundo-com-br.cdn.ampproject.org/c/s/www.tecmundo.com.br/amp/seguranca/223168-detran-sp-tem-vazamento-2-milhoes-dados.htm>. Acesso em 27 out. 2021.

compromisso com a proteção dos dados, diverenciar-se-ão de seus concorrentes e ganharão a preferência dos usuários[280].

A implementação de medidas que aumentam o grau de proteção dos dados tratados por determinada organização demandará desta investimentos de ordem econômica. O outro lado da moeda, contudo, seria ainda mais custoso: sem a mitigação de riscos, as bases de dados mantidas pelo agente ficam mais suscetíveis a vazamentos e acessos indevidos, tornando a organização sujeita a multas, que, a depender do seu porte econômico, podem causar grandes impactos à saúde financeira da empresa, e à degradação de sua imagem no mercado.

Além da implementação por entes privados para fins de segurança interna e conquista de confiança do público, a anonimização possui outros três principais usos: **(i)** quando a lei determina que a divulgação de determinados dados devem ser anonimizados; **(ii)** compartilhamento de dados com terceiros pela organização; **(iii)** publicização de bases de dados. No tocante ao primeiro, a LGPD, conforme já referido, pontua algumas formas de tratamento de dados em que se recomenda que os agentes apliquem a anonimização, seja por conta da natureza dos dados (dados sensíveis) ou pelo tipo de atividade desenvolvida (estatística e pesquisa). Em relação ao compartilhamento dos dados com terceiros, por meio da aplicação de técnicas de anonimização, a organização minimiza os riscos de danos à sua reputação em caso de publicação ou divulgação inadequada ou insegura de dados pessoais. Além disso, reduz eventuais perguntas e reclamações dos titulares acerca do compartilhamento dos dados[281]. Por fim, a publicização dos dados, como se pôde observar nos estudos relatados no Capítulo 2, é extremamente suscetível à reversão por combinação com outras bases de dados. Grande parte dos estudos recentemente desenvolvidos na área[282]

---

[280] SAAD, Andreia; HIUNES, Antonio. Ela, a LGPD, vista pelas empresas: uma proposta de visão prática – e otimista. In.: DONEDA, Danilo et. al. (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 27.

[281] UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. p. 09. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021.

[282] Nesse sentido: XU, Yabo et al. Anonymizing transaction databases for publication. In: ACM SIGKDD international conference on Knowledge discovery and data mining, ago. 2008, Las Vegas/NV. **Proceedings of the 14th [...]**. New York: Association for Computing Machinery – ACM, 2008. p. 767-775. Disponível em: <https://dl.acm.org/doi/pdf/10.1145/1401890.1401982?casa_token=r1BtefJR52UAAAAA:jCrRv_CcSPPXehhUzVRwShoDebLwTTyjH0ZJIQfaxM8tPaECO5N5I9Z82cyU08jN6W4W8eaDKEORwg>. Acesso em 27 out. 2021; e GUNAWAN, Dedi. A Data Anonymization Method to Mitigate Identity Attack in Transactional Database Publishing. In: International Conference on Information and Communication Technology (ICoICT), jun. 2020, Yogyakarta. **Proceedings of the 8th [...]**.Yogyakarta: IEEE, 2020. p. 1-6. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9166262&casa_token=YA_to2Fwd4cAAAAA:A8RKGpT7utsf98JrGwtoyePoFro9tb6YkIJ_CZO9zJeQ4DFJ2R05KP483yUZMsfRiZpDwhHxpCc&tag=1>. Acesso em 27 out. 2021.

tem como objetivo a diminuição dos riscos de reidentificação de bases de dados anonimizadas que são disponibilizadas ao público. A anonimização de dados para fins de compartilhamento público já foi, inclusive, objeto de decisão no STF. Na ocasião, discutia-se a possibilidade ou não da divulgação de dados de uma comunidade quilombola relacionados à pandemia da COVID-19. O Supremo proferiu decisão no sentido de que, considerando a situação excepcional gerada pela pandemia e a necessidade de proteção da saúde pública, é possível o compartilhamento de dados sensíveis dos quilombolas, mas destacou a necessidade da anonimização desses dados[283].

A anonimização dos dados é uma ferramenta de grande efetividade para fins de proteção dos titulares de dados e, consequentemente, de *compliance* com a LGPD. O seu uso conforme à Lei auxilia entes privados e o próprio poder público a consquistar maior confiança da sociedade através da garantia de que as salvaguardas legalmente exigidas estão em vigor e estão sendo cumpridas[284]. Para tanto, necessário o conhecimento dos conceitos previstos em Lei e dos critérios a serem atendidos para que a anonimização seja juridicamente reconhecida e alcance a sua finalidade de mitigação de riscos.

[283]BRASIL. Supremo Tribunal Federal. Petição 9.697/DF. Requerentes: Coordenação Nacional de Articulação das Comunidades Negras Rurais Quilombolas (CONAQ); Partido Socialista Brasileiro – PSB; Partido Socialismo e Liberdade - P-SOL; Rede Sustentabilidade; Partido dos Trabalhadores - PT. Requerido: União. Relator: Min. Edson Fachin. Brasília, 7 jul. 2021. Disponível em: <http://portal.stf.jus.br/processos/detalhe.asp?incidente=6187327>. Acesso em 20 out. 2021.

[284] UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. p. 09. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021.

5. CONSIDERAÇÕES FINAIS

A presente monografia buscou analisar como a Lei Geral de Proteção de Dados Pessoais – LGPD lida com o fenômeno da reversibilidade do processo de anonimização e de que forma este impacta a definição do escopo de aplicação da Lei. Com a entrada em vigor da LGPD, em 18 de setembro de 2020, iniciou-se uma grande movimentação do poder público e de entes privados para se adequar ao regime de proteção de dados pessoais. A implementação de PETs e outras tecnologias abrangidas pelo *Privacy by design* constituíram ferramentas importantes na corrida para a *compliance*. Neste cenário, a anonimização de dados tornou-se objeto de dúvidas atinentes à sua aplicação prática e sua eficácia. Com a finalidade de tornar mais claras algumas destas questões e apresentar a visão do Direito a respeito do tema, o estudo foi dividido em três capítulos.

No primeiro capítulo, fez-se uma síntese histórica sobre o surgimento de um regime de proteção de dados pessoais e sua relação com a ideia de privacidade para, depois, abordar-se sobre como o seu escopo de aplicação é delimitado pelo conceito de dado pessoal. Baseando-se em método adotado pelo Grupo de Trabalho do Artigo 29 (*Article 29 Data Protection Working Party* – WP29) no Parecer 04/2007, analisou-se, individualmente, cada um dos elementos que compõe tal definição, consoante as disposições do inciso I, artigo 5º da LGPD: **(i)** informação; **(ii)** relacionada; **(iii)** pessoa natural; **(iv)** identificada ou identificável. Objetivou-se verificar em quais circunstâncias um dado atrai para si a terminologia "pessoal" e em quais hipóteses ela é afastada. Em seguida, introduziu-se a questão da dualidade mutuamente excludente entre dados pessoais e dados anonimizados, e como esta se baseia, fundamentalmente, no grau de identificabilidade apresentado por determinada informação. Foi feita a distinção do conceito de dado anonimizado a outras duas "modalidades" de dados que são comumentemente confundidas àquele: os dados pseudonimizados e os dados criptografados.

No capítulo seguinte, abordou-se sobre a anonimização de dados pessoais e as oportunidades em que ela é citada em Lei. Posteriorente, dipôs-se a respeito da possibilidade de reversão do processo de anonimização e quais são os efeitos gerados no plano legal. Referiu-se que, na primeira geração de leis de proteção de dados, a anonimização era tida como um processo infalível, que sequer era quesitionado. Esta crença, denominada *robust anonymisation assumption*, começou a ser abandonada com a crescente publicação de estudos desenvolvidos na área da ciência da computação que comprovavam a falibilidade de vários procedimentos de anonimização antes considerados confiáveis. Relatou-se dois dos principais

trabalhos desenvolvidos na matéria: o caso do Censo dos EUA de 1990 e o caso do Netflix Prize. Revelou-se que, com o abandono da *robust anonymisation assumption*, o Direito passou a reconhecer o risco inerente de um dado anonimizado se transmudar em um dado pessoal, o que, todavia, criou uma redundância normativa naquelas regulações de proteção de dados que admitem que o titular dos dados seja meramente identificável. Conforme se demonstrou, as informações pessoais identificáveis poderiam configurar tanto um dado pessoal como um dado anonimizado, tornando a dualidade antes estabelecida algo destituído de sentido. Em seguida, apontou-se outro fator problemátio relativo à anonimização de dados, qual seja o *trade-off* utilidade dos dados *vs*. grau de anonimização. Demonstrou-se a existência de uma relação de proporcionalidade entre a identificabildiade de uma base de dados e a sua utilidade: quanto maior é a perda da identificabilidade de uma base de dados, menor são as possibilidades de uso. Foram relatados estudos desenvolvidos na área da ciência da computação em que diversas técnicas de anonimização de dados se mostraram falhas, permitindo a reidentificação dos titulares dos dados. Apresentou-se os principais grupos de técnicas de anonimização (randomização, generalização e supressão), abordando, ainda, sobre algumas submodalidades, tais como a adição de ruído, a permutação e o *k*-anonimato. Expô-se, por meio de exemplos práticos, como funciona a sua aplicação em bases de dados relacionais estruturadas em tabela e como a melhor maneira de alcançar níveis consideráveis de proteção se dá por meio da combinação de diferentes categorias de técnicas de anonimização.

Por fim, no terceiro capítulo, referiu-se que as interpretações jurídicas acerca da avaliação do nível de adequação de um processo de anonimização dividem-se em duas correntes: *risk-based approach* ou abordagem baseada no risco; e a *procedure-based approach* ou abordagem baseada nos procedimentos adotados. Tanto a LGPD quanto o GDPR adotam uma abordagem baseada no risco (*risk-based approach)* inerente de o processo de anonimização ser revertido, ou seja, as regulações não exigem que a anonimização seja um processo 100% eficaz para que um dado seja considerado anonimizado. Conforme explanado, para diferenciar dados pessoais de dados anonimizados, já que em ambas as categorias os titulares seriam identificáveis, passou-se a adotar o critério da razoabilidade. Abordou-se como a adoção de um conceito jurídico indeterminado, tal como a razoabilidade, pode permitir que a Lei não caia em defasagem, já que relacionada em grande escala com a tecnologia. Foi apontado que, apesar de permitir a neutralidade tecnológica, a razoabilidade, assim como a maioria dos conceitos jurídicos indeterminados, abrem espaço para a discricionariedade quando da interpretação normativa. Assim, demonstrou-se que para reduzir

a incerteza que tal critério poderia trazer, a Lei indica parâmetros que podem ser divididos em dois eixos de análise: o eixo subjetivo (meios próprios) e o eixo objetivo (custo, tempo e estado da arte da tecnologia). Fazendo o uso de uma representação esquemática, foram enumerados os passos que devem ser seguidos quando da realização do "teste da razoabilidade", a fim de determinar se o dado é categorizado como dado anonimizado ou dado pessoal. Foram expostas algumas peculiaridades de cada um dos eixos de análise, oportunidade em que foram apontadas lacunas existentes na Lei brasileira, como, por exemplo, a consideração ou não de meios ilícitos que possibilitem a reversão da anonimização. Para concluir, dispôs-se sobre como a anonimização de dados, ainda que não reduzindo a zero a possibilidade de reidentificação do titular, pode ser uma grande aliada na mitigação dos riscos advindos do tratamento de dados pessoais, trazendo maior segurança e confiabilidade quando do desenvolvimento de tal atividade.

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, Bethania de Araujo et al. Preservação da privacidade no enfrentamento da COVID-19: dados pessoais e a pandemia global. **Ciência & Saúde Coletiva [online]**, v. 25, p. 2487-2492, jun. 2020. Disponível em: <https://doi.org/10.1590/1413-81232020256.1.11792020>. Acesso em 24 out. 2021.

ARANHA, Diego F. O que é criptografia fim a fim e o que devemos fazer a respeito? In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. p. RB-1.1- RB-1.5. *E-book*.

ARBIX, Daniel. A importância da privacidade por *design* e por *default* (*privacy by design and by default*). In.: DONEDA, Danilo et. al. (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 55-63.

BARACHO, José Alfredo de Oliveira. Teoria geral dos conceitos legais indeterminados. **Themis**, v. 2, n. 2, p. 61-78, 1999.

BENNET, Colin J.; GRANT, Rebecca. **Visions Of Privacy: Policy Choices For The Digital Age**. Toronto: University of Toronto Press, Scholarly Publishing Division, 1999.

BRASIL. **Lei nº 5.534, de 14 de novembro de 1968**. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l5534.htm>. Acesso em 20 out. 2021.

BRASIL. **Decreto-lei nº 4.657, de 4 de setembro de 1942**. Lei de Introdução às Normas do Direito Brasileiro. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto-lei/del4657compilado.htm>. Acesso em 20 out. 2021.

BRASIL. **Decreto nº 10.153, de 3 de dezembro de 2019**. Disponível em: <https://www.in.gov.br/web/dou/-/decreto-n-10.153-de-3-de-dezembro-de-2019-231274119>. Acesso em 17 out. 2021.

BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002**. Código Civil. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm>. Acesso em 01 set. 2021.

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011.** Lei de Acesso à Informação (LAI). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em 01 set. 2021.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em 01 set. 2021.

BRASIL. Governo Federal. O que são dados sensíveis, de acordo com a LGPD. Disponível em: <https://www.serpro.gov.br/lgpd/menu/protecao-de-dados/dados-sensiveis-lgpd>. Acesso em 21 set. 2021.

BRASIL. Governo Federal. Perguntas e Respostas. Disponível em: <https://www.gov.br/ouvidorias/pt-br/central-de-conteudos/perguntas-frequentes-2019#resp1>. Acesso em 16 out. 2021.

BRASIL. Supremo Tribunal Federal. ADIs no 6.387, 6.388, 6.389, 6.390 e 6.393. Relatora: Ministra Rosa Weber. 7 de maio de 2020. Disponível em: <http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895165> Acesso em 21 set. 2021.

BRASIL. Supremo Tribunal Federal. Arguição de Descumprimento de Preceito Fundamental 403/SE. **Voto do Relator Min. Luiz Edson Fachin**. Disponível em: <https://www.conjur.com.br/dl/fachin-suspensao-whatsapp-decisao.pdf>. Acesso em 24 set. 2021.

BRASIL. Supremo Tribunal Federal. Petição 9.697/DF. Requerentes: Coordenação Nacional de Articulação das Comunidades Negras Rurais Quilombolas (CONAQ); Partido Socialista Brasileiro – PSB; Partido Socialismo e Liberdade - P-SOL; Rede Sustentabilidade; Partido dos Trabalhadores - PT. Requerido: União. Relator: Min. Edson Fachin. Brasília, 7 jul. 2021. Disponível em: <http://portal.stf.jus.br/processos/detalhe.asp?incidente=6187327>. Acesso em 20 out. 2021.

BIONI, Bruno Ricardo. **Calibrando o filtro da razoabilidade: critérios objetivos e subjetivos como fatores de uma análise de risco**. 14 nov. 2019. Disponível em: <http://genjuridico.com.br/2019/11/14/filtro-da-razoabilidade-criterios/>. Acesso em 27 out. 2021.

BIONI, Bruno Ricardo. Compreendendo o conceito de anonimização e dado anonimizado. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 39-54.

BIONI, Bruno Ricardo. **Proteção de dados pessoais**: a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

BIONI, Bruno Ricardo. **Xeque-Mate**: o tripé da proteção de dados pessoais no jogo de xadrez das iniciativas legislativas no Brasil. Privacidade e Vigilância. São Paulo: GPoPAI/USP, 2015.

BUCHNER, Benedikt; KÜHLING, Jürgen (org.). **Datenschutz-Grundverordnung BDSG** – Kommentar. 3. ed. München: C.H.Beck, 2020. *E-book*.

CADWALLADR, Carole; GRAHAM-HARRISON, Emma. **Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach**. 17 mar. 2018. Disponível em: <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>. Acesso em 25 set. 2021.

CHAMOUX, Jean Pierre. Data Protection in Europe - The Problem of the Physical Person and the Legal Person. **Journal of Media Law and Practice,** v. 2, n. 1, p. 70-83, 1981.

COUNCIL OF EUROPE. **Contact Tracing Apps**. 10 jun. 2020. Disponível em: <https://www.coe.int/en/web/data-protection/contact-tracing-apps>. Acesso em 24 out. 2021.

CORDEIRO, A. Barreto Menezes. **Direito da Proteção de dados**: à luz da RGPD e da Lei n.º 58/2019. Coimbra: Al Medina, 2020.

CORRÊA, Adriana Espíndola; DA LUZ, Pedro Henrique Machado. **A exceção na proteção de dados pessoais durante a Covid-19 - parte 1**. 22 mai. 2020. Disponível em: <https://www.conjur.com.br/2020-mai-22/direito-civil-atual-excecao-protecao-dados-pessoais-durante-covid-19>. Acesso em 24 out. 2021.

CORRÊA, Adriana Espíndola; DA LUZ, Pedro Henrique Machado. **A exceção na proteção de dados pessoais durante a Covid-19 – parte 2**. 23 mai. 2020. Disponível em: <https://www.conjur.com.br/2020-mai-23/direito-civil-atual-excecao-protecao-dados-pessoais-durante-covid-19>. Acesso em 24 out. 2021.

DATAGUIDANCE BY ONETRUST; BAPTISTA LUZ ADVOGADOS**. Comparing privacy laws: GDPR v. LGPD**. 2019. Disponível em: <https://baptistaluz.com.br/wp-content/uploads/2019/05/DataGuidance-GPDR-LGPD-For-Print.pdf>. Acesso em 28 out. 2021.

DEUTSCHLAND. BVerfGE 65, 1. Bundesverfassungsgericht, Karlsruhe, 1983. Disponível em: <https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/1983/12/rs19831215_1bvr020983.html>. Acesso em 20 set. 2021.

DOLIN, Ron A. Search Query Privacy: The Problem of Anonymization. **Hastings Science and Technology Law Journal**, v. 2, n. 2, p. 137-182, mai. 2010.

DONEDA, Danilo. **A proteção de dados em tempos de coronavírus**. 25 mar. 2020. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/a-protecao-de-dados-em-tempos-de-coronavirus-25032020>. Acesso em 21 out. 2021.

DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos da formação da Lei Geral de Proteção de Dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019.

DONEDA, Danilo. Panorama histórico da proteção de dados pessoais. In: BIONI, Bruno et al. (coord.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. p. 3-20.

DONEDA, Danilo. A proteção dos dados pessoais como um direito fundamental. **Espaço Jurídico**, v. 12, n. 2, , p. 91-108, jul./dez. 2011.

DONEDA, Danilo; MACHADO, Diego. Proteção de dados pessoais e Criptografia: Tecnologias Criptográficas Entre Anonimização E Pseudonimização de Dados. In: DONEDA, Danilo; MACHADO, Diego (coords.). **A Criptografia no Direito Brasileiro**. 1. ed. São Paulo: Thomson Reuters Brasil, 2019. p. RB-8.1- RB-8.5. *E-book*.

DOURIEZ, Marie (et. al.). Anonymizing NYC Taxi Data: Does It Matter?. In: IEEE International Conference on Data Science and Advanced Analytics (DSAA), out. 2016,

Montréal/QC. **Proceedings of 2016 […]**. Montréal: IEEE, 2016. p. 140-148. Disponível em: <https://ieeexplore.ieee.org/document/7796899>. Acesso em 27 out. 2021.

EL EMAM, Khaled; HINTZE, Mike. **Does anonymization or de-identification require consent under the GDPR?**. 29 jan. 2019. Disponível em: <https://iapp.org/news/a/does-anonymization-or-de-identification-require-consent-under-the-gdpr/>. Acesso em 20 out. 2021.

EUROPEAN COMMISSION. **The Article 29 Working Party ceased to exist es of 25 May 2018**. Disponível em: <https://ec.europa.eu/newsroom/article29/items/629492/en>. Acesso em 07 set. 2021.

EUROPEAN UNION. Collaboration in Research and Methodology for Official Statistics - CROS. **Indirect identification**. Disponível em: <https://ec.europa.eu/eurostat/cros/content/indirect-identification_en>. Acesso em 16 out. 2021.

EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 4/2007 on the concept of personal data**. Brussels, 20 jun. 2007. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>. Acesso em 21 set. 2021.

EUROPEAN UNION. Article 29 Data Protection Working Party. **Opinion 5/2014 on Anonymisation Techniques. Brussels**, 10 abr. 2014. Disponível em: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. Acesso em 21 set. 2021.

EUROPEAN UNION. Regulation (EU) 2016/679. **General Data Protection Regulation – GDPR**. Disponível em: <https://gdpr-info.eu/>. Acesso em 25 set. 2021.

EUROPEAN UNION. Recital 14. Regulation (EU) 2016/679. **General Data Protection Regulation – GDPR**. Disponível em: <https://www.privacy-regulation.eu/en/recital-14-GDPR.htm>. Acesso em 25 set. 2021.

EUROPEAN UNION. Recital 26. Regulation (EU) 2016/679. **General Data Protection Regulation – GDPR**. Disponível em: <https://gdpr-text.com/pt/read/recital-26/>. Acesso em 21 set. 2021.

EUROPEAN UNION AGENCY FOR NETWORK AND INFORMATION SECURITY - ENISA. **Privacy by design in big data: An overview of privacy enhancing technologies in the era of big data analytics**. Heraklion, dez. 2015. Disponível em: <https://www.enisa.europa.eu/publications/big-data-protection>. Acesso em 20 out. 2021.

FINCK, Michèle; PALLAS, Frank. They who must not be identified - distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, v. 10, n. 1, p. 11-36, 2020.

FLORIDI, Luciano; TAYLOR, Linnet; VAN DER SLOOT, Bart. Introduction: A New Perspective on Privacy. In: FLORIDI, Luciano; TAYLOR, Linnet; VAN DER SLOOT, Bart

(coord.). **Group Privacy**: New Challenges of Data Technologies. Dordrecht: Springer International Publishing, 2017. p. 1-12. *E-book*.

FONSECA, Gabriel Campos Soares da; MENDES, Laura Schertel. STF reconhece direito fundamental à proteção de dados - Comentários sobre o referendo da Medida Cautelar nas ADIs 6387, 6388, 6389, 6390 e 6393. **Revista de Direito do Consumidor**, v. 130, p. 471-478, 2020.

FUSTER, Gloria González. **The Emergence of Personal Data Protection as a Fundamental Right of the EU**. Bruxelas: Springer Science & Business, 2014.

GOOGLE. **Primeiros passos com arquivos criptografados no Drive, no Documentos, no Planilhas e no Apresentações.** Disponível em: <https://support.google.com/docs/answer/10519333?hl=pt-BR>. Acesso em 21 out. 2021.

GRACE, Paul et al. **D4.3 – Guidelines for data anonymization report**. Out. 2016. Disponível em: <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ade9ca8a&appId=PPGMS>. Acesso em 05 nov. 2021.

GRESHAM, Joshua. **Is encrypted data personal data under the GDPR?** 6 mar. 2019. Disponível em: <https://iapp.org/news/a/is-encrypted-data-personal-data-under-the-gdpr/>. Acesso em 21 out. 2021.

GUNAWAN, Dedi. A Data Anonymization Method to Mitigate Identity Attack in Transactional Database Publishing. In: International Conference on Information and Communication Technology (ICoICT), jun. 2020, Yogyakarta. **Proceedings of the 8th [...]**.Yogyakarta: IEEE, 2020. p. 1-6. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9166262&casa_token=YA_to2Fwd4cAAAAA:A8RKGpT7utsf98JrGwtoyePoFro9tb6YkIJ_CZO9zJeQ4DFJ2R05KP483yUZMsfRiZpDwhHxpCc&tag=1>. Acesso em 27 out. 2021.

HOFFMANN, Bruno. **Falha da Enel expõe dados de milhões de clientes**. 04 fev. 2020. Disponível em: <https://www.gazetasp.com.br/estado/2020/02/1061629-falha-da-enel-expoe-dados-de-milhoes-de-clientes.html>. Acesso em 27 out. 2021.

JR., Sérgio Alves. Fechando um ciclo: do término do tratamento de dados pessoais (arts. 15 e 16 da LGPD). In: BIONI, Bruno et al (coord.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. p. 227-241.

JÚNIOR, Odélio Porto. **Anonimização e Pseudonimização: conceitos e diferenças na LGPD**. 25 mai. 2019. Disponível em: <https://baptistaluz.com.br/espacostartup/anonimizacao-e-pseudonimizacao-conceitos-e-diferencas-na-lgpd/#_ftn4>. Acesso em 04 out. 2021.

KNAPP, Gerry; OROOJI, Marmar. Improving Suppression to Reduce Disclosure Risk and Enhance Data Utility. In: IISE Annual Conference, mai. 2018, Orlando/FL. **Proceedings of the 2018** [...]. Orlando: Institute of Industrial & Systems Engineers (IISE), 2018. p. 1415-1420. Disponível em: <https://arxiv.org/pdf/1901.00716.pdf>. Acesso em 21 set. 2021.

LEOCÁDIO, Thais; PIMENTA, Guilherme. **Dados pessoais e jornada de trabalho dos professores da rede estadual vazam na internet**. 14 mai. 2021. Disponível em: < https://g1.globo.com/mg/minas-gerais/noticia/2021/05/14/dados-pessoais-e-jornada-de-trabalho-dos-professores-da-rede-estadual-vazam-na-internet.ghtml>. Acesso em 27 out. 2021.

LIU, Junqiang. Privacy Preserving Data Publishing: Current Status and New Directions. **Information Technology Journal**, v. 11, n. 1, p. 1-8, 2012.

MAGALHÃES, Rodrigo Almeida; DIVINO, Sthéfano Bruno Santos. A proteção de dados da pessoa jurídica e a Lei 13.709/2018: reflexões à luz dos direitos da personalidade. **Scientia Iuris**, v. 23, n. 2, p. 74-90, jul. 2019.

MARTINS-COSTA, Judith. **A boa-fé no direito privado**: critérios para a sua aplicação. 2. ed. São Paulo: Saraiva Educação, 2018. *E-book*.

MAYER-SCHÖNBERGER Viktor. Generational Development of Data Protection in Europe. In: AGRE, Philip E.; ROTENBERG, Marc (orgs.). **Technology and Privacy:** The New Landscape. Cambridge: The MIT Press, 1997. p. 219-241.

MENDES, Laura Schertel. **A encruzilhada da proteção de dados no Brasil e o caso do IBGE**. 23 abr. 2020. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/a-encruzilhada-da-protecao-de-dados-no-brasil-e-o-caso-do-ibge-23042020#sdfootnote2sym>. Acesso em 21 out. 2021.

MENKE, Fabiano. **As origens alemãs e o significado da autodeterminação informativa**. 30 out. 2020. Disponível em: <https://www.migalhas.com.br/coluna/migalhas-de-protecao-de-dados/335735/as-origens-alemas-e-o-significado-da-autodeterminacao-informativa>. Acesso em 24 out. 2021.

MERKLE, Ralph C. A fast software one-way hash function. **Journal of Cryptology**, v. 3, n. 1, p. 43-58, jul. 1990.

NALDI, Maurizio. **Anonymization Systems and Utility**. IPEN - Workshop Rome, 12 jun. 2019. Disponível em: <https://edps.europa.eu/sites/default/files/publication/12-06-19_maurizio-naldi_anonymization-systems-and-utility_en_0.pdf>. Acesso em 22 out. 2021.

NARAYANAN, Arvind; SHMATIKOV, Vitaly. How To Break Anonymity of the Netflix Prize Dataset. **ArXiv e-prints**, out. 2006. Disponível em: <https://arxiv.org/abs/cs/0610105#>. Acesso em 25 out. 2021.

Netflix annual revenue hits US$20 billion. 1 vídeo (4 min). Publicado por BNN – Bloomberg. 22 jan. 2020. Disponível em: <https://www.bnnbloomberg.ca/technology/video/netflix-annual-revenue-hits-us-20-billion~1881767>. Acesso em 25 out. 2021.

OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. **UCLA Law Review**, v. 57, n. 6, p. 1701-1778, 2010.

PAAR, Christof; PELZL, Jan. **Understanding cryptography**: a textbook for students and practitioners. Londres: Springer, 2010.

PAGALLO, Ugo. The Group, the Private, and the Individual: A New Level of Data Protection? In: FLORIDI, Luciano; TAYLOR, Linnet; VAN DER SLOOT, Bart (coord.). **Group Privacy**: New Challenges of Data Technologies. Dordrecht: Springer International Publishing, 2017. p. 159-173. *E-book*.

PANDURANGAN, Vijay. **On taxis and rainbows: Lessons from NYC's improperly anonymized taxi logs**. Medium. 21 jun. 2014. Disponível em: <https://tech.vijayp.ca/of-taxis-and-rainbows-f6bc289679a1>. Acesso em 27 0ut. 2021.

PAYÃO, Felipe. **Detran SP tem vazamento de 2 milhões de dados**. 16 ag. 2021. Disponível em: <https://www-tecmundo-com-br.cdn.ampproject.org/c/s/www.tecmundo.com.br/amp/seguranca/223168-detran-sp-tem-vazamento-2-milhoes-dados.htm>. Acesso em 27 out. 2021.

PERSONAL DATA PROTECTION COMMISSION OF SINGAPORE – PDPC. **Guide to basic data anonymisation techniques**. Singapore, 25 jan. 2018. Disponível em: <https://iapp.org/resources/article/guide-to-basic-data-anonymization-techniques/>. Acesso em 20 out. 2021.

PORMEISTER, Kart; DROZDZOWSKI, Lukasz. Protecting the Genetic Data of Unborn Children: A Critical Analysis. **European Data Protection Law Review (EDPL)**, v. 4, n. 1, p. 53-64, 2018.

PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. **Law, Innovation and Technology**, v. 10, n. 1, p. 41-81, 2018.

REALE, Miguel. **Lições preliminares de direito**. 27. ed. São Paulo: Saraiva, 2002.

ROZAS, Luiza Barros. Conceitos jurídicos indeterminados e discricionariedade administrativa. **Cadernos Jurídicos da Escola Paulista de Magistratura**, v. 20, n. 47, p. 191-201, 2019.

SAAD, Andreia; HIUNES, Antonio. Ela, a LGPD, vista pelas empresas: uma proposta de visão prática – e otimista. In.: DONEDA, Danilo; MENDES, Laura Schertel; CUEVA, Ricardo Villas Bôas (coord.). **Lei Geral de Proteção de Dados (Lei nº 13.709/2018) - A caminho da efetividade**: contribuições para a implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020. p. 17-28.

SAMARATI, Pierangela; SWEENEY, Latanya. Protecting Privacy when Disclosing Information: *k*-Anonymity and Its Enforcement through Generalization and Suppression. **Tech Report**, 1998. Disponível em: <http://www.csl.sri.com/papers/sritr-98-04/>. Acesso em 20 out. 2021.

SARLET, Ingo Wolfgang. Fundamentos constitucionais: o direito fundamental à proteção de dados. In: BIONI, Bruno et al. (coord.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. p. 21-59.

SAYAS, Samson Yoseph. The role of anonymisation and pseudonymisation under the EU data privacy rules: beyond the 'all or nothing' approach. **European Journal of Law and Technology**, v. 6, n. 2, p. 1-28, 2015.

SÃO PAULO. 3ª Vara de Fazenda Pública. Ação Popular 1020192-74.2020.8.26.0053. Requerente: Mauricio Roberto Giosa. Requerido: Fazenda Pública do Estado de São Paulo. Julgador: Juiz Luis Manuel Fonseca Pires. São Paulo, 22 jun. 2020. Disponível em: <https://esaj.tjsp.jus.br/cpopg/show.do?processo.codigo=1H000HDVE0000&processo.foro=53&processo.numero=1020192-74.2020.8.26.0053&uuidCaptcha=sajcaptcha_bfe33b253a5443ef808ae42740df88c3>. Acesso em 24 out. 2021.

SCHWARTZ, Paul M; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. **New York University Law Review**, v. 86, n. 6, p. 1814-1894, 2011.

SPINDLER, Gerald; SCHMECHEL, Philipp. Personal Data and Encryption in the European General Data Protection Regulation. **Journal of Intellectual Property, Information Technology and Electronic Commerce Law**, v. 7, n. 2, p. 163-177, set. 2016.

SWEENEY, Latanya. Achieving k-anonymity privacy protection using generalization and suppression. **International Journal of Uncertainty, Puzziness and Knowledge-Based Systems**, v. 10, n. 5, p. 571–588, 2002.

SWEENEY, Latanya. *k*-Anonymity: a model for protecting privacy. **International Journal on Uncertainty, Fuzziness and Knowledge-based Systems**, v. 10, n. 5, p. 557-570, 2002.

SWEENEY, Latanya. Simple demographics often identify people uniquely. **Health (San Francisco)**, v. 671, n. 2000, p. 1-34, 2000.

UNCTAD. **Data Protection And Privacy Legislation Worldwide**. 2020. Disponível em: <https://unctad.org/page/data-protection-and-privacy-legislation-worldwide>. Acesso em 09 set. 2021.

UNIÃO EUROPEIA. TJUE. Processo n. C-582/14 - Breyer. Autor: Patrick Breyer. Réu: Bundesrepublik Deutschland. 19 out. 2016. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A62014CJ0582>. Acesso em 27 out. 2021.

UK INFORMATION COMMISSIONER'S OFFICE – ICO. **Anonymisation: managing data protection risk code of practice**. Wilmslow, nov. 2012. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>. Acesso em 17 abr. 2021.

UK INFORMATION COMMISSIONER'S OFFICE – ICO. **What is automated individual decision-making and profiling?**. Disponível em: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/>. Acesso em 20 out. 2021.

VAN DER SLOOT, Bart; BROEDERS, Dennis; SCHRIJVERS, Erik (org.). **Exploring the boundaries of Big Data**. Amsterdam: WRR, 2016.

VITAL, Danilo. **Rosa Weber atende OAB e suspende MP do compartilhamento de dados**. 24 abr. 2020. Disponível em: < https://www.conjur.com.br/2020-abr-24/rosa-atende-oab-suspende-mp-compartilhamento-dados>. Acesso em 20 set. 2021.

VOKINGER, Kerstin N.; STEKHOVEN, Daniel J.; KRAUTHAMMER, Michael. Lost in Anonymization — A Data Anonymization Reference Classification Merging Legal and Technical Considerations. **The Journal of Law, Medicine & Ethics**, v. 48, n. 1, p. 228–231, 2020.

WHATSAPP. **Sobre a criptografia de ponta a ponta.** Disponível em: <https://faq.whatsapp.com/general/security-and-privacy/end-to-end-encryption/?lang=pt_br> Acesso em 21 out. 2021.

Webinário: Anonimização e pseudonimização de dados pessoais - conversas sobre a LGPD e o papel da Ouvidoria. 1 vídeo (162 min). Publicado pelo canal Controladoria-Geral da União – CGU. Disponível em: <https://www.youtube.com/watch?v=8AVj0wmzFRs>. Acesso em: 29 jun. 2021.

WIGMORE, Ivy. **Unique Identifier (UID)**. Disponível em: <https://internetofthingsagenda.techtarget.com/definition/unique-identifier-UID>. Acesso em 11 out. 2021.

XU, Yabo et al. Anonymizing transaction databases for publication. In: 14th ACM SIGKDD international conference on Knowledge discovery and data mining, ago. 2008, Las Vegas/NV. **Proceedings of the 14th [...]**. New York: Association for Computing Machinery – ACM, 2008. p. 767-775. Disponível em: <https://dl.acm.org/doi/pdf/10.1145/1401890.1401982?casa_token=r1BtefJR52UAAAAA:jCrRv_CcSPPXehhUzVRwShoDebLwTTyjH0ZJIQfaxM8tPaECO5N5I9Z82cyU08jN6W4W8eaDKEORwg>. Acesso em 27 out. 2021.