ESTUDO TÉCNICO SOBRE A ANONIMIZAÇÃO DE DADOS NA LGPD: ANÁLISE JURÍDICA

VERSÃO 1.0

BRASÍLIA/DF
NOVEMBRO DE 2023

Autoridade Nacional de Proteção de Dados

Diretor-Presidente
Waldemar Gonçalves Ortunho Júnior

Diretores
Arthur Pereira Sabbat
Joacil Basílio Rael
Miriam Wimmer

Equipe de Elaboração
Marcelo Santiago Guedes – Coordenador-Geral de Tecnologia e Pesquisa (CGTP)
Diego Carvalho Machado – Especialista (CGTP)
Albert França Josuá Costa – Especialista (CGTP)

SUMÁRIO

INTRODUÇÃO

Previstos respectivamente no art. 5º, incisos III e XI, da Lei Geral de Proteção de Dados Pessoais (LGPD), os conceitos de "dado anonimizado" e "anonimização" estão entre as noções fundamentais de maior relevância para o sistema brasileiro de proteção de dados, visto que ambos possuem como principal repercussão jurídica o afastamento de obrigações e encargos regulatórios impostos pelo regime geral estabelecido na LGPD para agente de tratamento de dados pessoais. Além das consequências jurídicas vinculadas aos referidos conceitos, que convergem no que a lei denomina de "processo de anonimização"[1], o tema ganha em importância devido aos seus fundamentos computacionais, que ensejam muitos debates na comunidade acadêmica a respeito dos limites da anonimização de dados e de suas técnicas na atualidade.

Tendo em vista se tratar de assunto fundamental para os regimes de proteção de dados pessoais, como o estabelecido pela LGPD, e a complexidade do debate técnico e jurídico, autoridades de proteção de dados mundo afora se preocuparam com a elaboração de estudos, guias e documentos relativos à anonimização de dados. Este estudo da Autoridade Nacional de Proteção de Dados (ANPD) visa a compreender os fundamentos jurídico-normativos do processo de anonimização de dados na sistemática da LGPD no ordenamento jurídico brasileiro. A fixação do entendimento e interpretação institucional sobre os fundamentos do processo de anonimização lançará as bases para futuras contribuições e orientações da ANPD – de cunho técnico e computacional, por exemplo – a todos os atores envolvidos no ecossistema nacional de proteção de dados pessoais.

O estudo está estruturado em quatro capítulos. Na primeira parte, é delimitada a metodologia em que o trabalho está ancorado. Em seguida, desenvolve-se uma análise jurídica do processo de anonimização, com a delimitação dos conceitos básicos e correlatos a dados anonimizados e anonimização. No capítulo 3, por sua vez, a concretização dos princípios de proteção de dados pertinentes à anonimização é abordada, assim como a definição das linhas gerais do modelo ou abordagem baseada em riscos de (re)identificação, constante da LGPD. Por fim, o capítulo 4 apresenta de forma

[1] LGPD, art. 12.

consolidada as principais conclusões e recomendações a respeito do processo de anonimização de dados na sistemática da LGPD.

1. METODOLOGIA

Trata-se de pesquisa teórica orientada pelos tipos metodológicos jurídico-compreensivo e jurídico-comparativo.[2] Tal estruturação metodológica é justificada pela necessidade de analisar e compreender os fundamentos normativos do processo de anonimização de dados na disciplina jurídica da LGPD no sistema brasileiro. Para isso, é igualmente necessário o confronto com outros conceitos jurídicos afins ou correlacionados igualmente inscritos na LGPD.

Entretanto, além dessa comparação na perspectiva interna da LGPD e do sistema jurídico do Brasil, realiza-se uma comparação entre conceitos e regimes jurídicos do sistema brasileiro e do ordenamento comunitário da União Europeia (UE) conforme o método funcionalista, inclusive devido à influência do Regulamento n. 2016/679 da UE sobre a vigente lei brasileira de proteção de dados[3]. Entende-se que a função desempenhada por conceitos como dado pessoal e dado não pessoal no Regulamento n. 2016/679 e no direito vinculante da UE é similar àqueles de dado pessoal, dado não pessoal e dado anonimizado estabelecidos na LGPD e aplicados no direito brasileiro.

Vale ressaltar, ainda, que o uso desse método comparativo não se atém apenas aos aspectos legislativos e doutrinários, mas também a outros elementos da prática em proteção de dados pessoais (decisões judiciais, manifestações de autoridades de proteção de dados, técnicas computacionais etc.), e com a consciência das suas limitações à luz da literatura contemporânea[4].

---

[2] GUSTIN, Miracy B. S.; DIAS, Maria Tereza Fonseca; NICÁCIO, Camila Silva. **(Re)pensando a pesquisa jurídica***:* teoria e prática. 5. ed. rev., amp. e atual. São Paulo: Almedina, 2020. p. 80-85.

[3] DONEDA, Danilo; MENDES, Laura Schertel. Um perfil da nova Lei Geral de Proteção de Dados brasileira. *In*: BELLI, Luca; CAVALLI, Olga (Orgs.). **Governança e Regulações da Internet na América Latina**. Rio de Janeiro: FGV Direito Rio, 2018. p. 309. A influência se estende também a outros países latino-americanos: NOUGRÈRES, Ana Brian. Data Protection and Enforcement in Latin America and in Uruguay. *In*: Wright David; DE HERT, Paul (Orgs.). **Enforcing Privacy:** Regulatory, Legal and Technological Approaches. Cham: Springer, 2016. p. 146, 148.

[4] Cf. MICHAELS, Ralf. The Functional Method of Comparative Law. *In*: REIMANN, M.; ZIMMERMANN, R. (Orgs.). **The Oxford Handbook of Comparative Law**. 2. ed. Oxford: Oxford Univesity Press, 2019. p. 339–382.

3. DEFINIÇÃO E CONCEITOS CORRELATOS

O conceito de **anonimização**, previsto pelo art. 5º, XI, da LGPD, é definido 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 legislador também estabeleceu a noção conceitual de **dado anonimizado**, que resulta da implementação da anonimização por agente de tratamento. Nos termos da LGPD, tal dado é considerado o "dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento"[5].

A partir da análise do art. 12, *caput*[6], da lei de proteção de dados brasileira, porém, depreende-se que a utilização de meios técnicos na anonimização de dados consiste, na verdade, em um conjunto de atos ou medidas entre si relacionadas que fazem parte de um **processo**[7]. Assim, a anonimização de dados se desenvolve em uma série de etapas que se inicia com o processamento de dado de caráter pessoal e tem o objetivo de, com a aplicação de técnicas variadas, desassociar identificadores do dado em seu estado originário ou bruto.[8]

O objetivo da anonimização é afetar os **identificadores** presentes em um dado ou conjunto de dados, porque esses são os elementos informativos que "mantém relação

---

[5] LGPD, art. 5º, III.

[6] "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."

[7] A noção de processo empregada neste estudo mais se aproxima de um tipo de procedimento, isto é, uma sucessão de atos encadeados entre si orientados para o alcance de um fim. Trata-se, notadamente, de um procedimento técnico de tratamento de dados. Sobre o termo "processo" no campo jurídico: cf. COUTO E SILVA, Clóvis V. do. **A obrigação como processo**. Rio de Janeiro: Editora FGV, 2006. p. 20, 63; DIDIER JR., Fredie. **Curso de direito processual civil**. 17. ed. rev., atual. e amp. Salvador: JusPodivm, 2015. v. 1. p. 30-31.

[8] A conceptualização da anonimização como **processo** tem sido adotada por várias autoridades de proteção de dados, havendo aquelas com estudos já publicados sobre o tema: cf. GRUPO DE TRABALHO DE PROTEÇÃO DE DADOS DO ARTIGO 29. **Parecer 05/2014 sobre técnicas de anonimização** Bruxelas: [s. n.], 2014. p. 5; INFORMATION COMMISSIONER'S OFFICE. **Anonymisation:** managing data protection risk code of practice. Wilmslow: ICO, 2012. p. 28; AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. **Orientaciones y garantias em los procedimentos de anonimización de datos personales**. [S.l.]: AEPD, 2016. p. 5; DATA PROTECTION COMMISSION. **Guidance on Anonymisation and Pseudonymisation.** [S.l.]: DPC, 2019. p. 14; AGENCIA DE GOBIERNO ELECTRÓNICO Y SOCIEDAD DE LA INFORMACIÓN Y DEL CONOCIMIENTO. **Guía sobre Anonimización de Datos**. [S.l]: AGESIC, 2020. p. 4; PERSONAL DATA PROTECTION COMMISSION. **Guide to basic anonymisation**. Singapore: SG Digital-PDPC, 2022. p. 13 et seq.

particularmente privilegiada e próxima com certo indivíduo"[9]. Os referidos identificadores podem ser, por sua vez, **diretos** ou **indiretos**. Identificador direto é o dado que permite identificar unicamente uma pessoa natural, sem a necessidade de combiná-lo com dados de outras fontes. O típico **identificador direto** de um titular de dados é o seu nome completo. Outro exemplo é o número de inscrição no Cadastro de Pessoas Físicas (CPF), que, a partir da entrada em vigor da Lei nº 14.534/2023, será considerado "número único e suficiente para identificação do cidadão nos bancos de dados de serviços públicos"[10].

Já **identificador indireto** é considerado o dado que por si só não tem a capacidade de identificar alguém, mas pode ser agregado e vinculado a dados auxiliares para identificar uma pessoa natural, a exemplo da nacionalidade, a idade, a raça, o CEP da residência, as características fenotípicas, ou o endereço de IP que podem ser necessários para distinguir (*single out*) alguém. Também conhecidos como "quase-identificadores"[11], os identificadores indiretos se relacionam ao "fenômeno das 'combinações únicas'"[12], isto é, tendo em vista que os atributos dos quase-identificadores variam de pessoa a pessoa que a combinação pode se tornar suficientemente singular a um único indivíduo. Por exemplo, no ano de 2000 a professora Latanya Sweeney publicou estudo demonstrando que 87% da população dos Estados Unidos da América[13] possuía características provavelmente únicas com base apenas no CEP de cinco dígitos (*5-digit ZIP code*), gênero e data de nascimento.[14]

Outro conceito que aparece com alguma proximidade à anonimização de dados é o de **pseudonimização**. Pode se dizer que em ambos os casos ocorre, de alguma forma,

---

[9] GRUPO DE TRABALHO DE PROTEÇÃO DE DADOS DO ARTIGO 29. **Parecer 4/2007 sobre o conceito de dados pessoais**. Bruxelas: [s. n.], 2007. p. 12.

[10] Art. 1º, *caput*, da referida lei. BRASIL. **Lei nº 14.534, de 11 de janeiro de 2023.** Altera as Leis nºs 7.116, de 29 de agosto de 1983, 9.454, de 7 de abril de 1997, 13.444, de 11 de maio de 2017, e 13.460, de 26 de junho de 2017, para adotar número único para os documentos que especifica e para estabelecer o Cadastro de Pessoas Físicas (CPF) como número suficiente para identificação do cidadão nos bancos de dados de serviços públicos. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2023/lei/l14534.htm. Acesso em: 09 mai. 2023.

[11] Cf. VIMERCATI, Sabrina de C., FORESTI, Sara. Quasi-Identifier. *In*: VAN TILBORG, Henk C. A.; JAJODIA, Sushil (Orgs.). **Encyclopedia of Cryptography and Security**. Springer: Boston, 2011.

[12] GRUPO DE TRABALHO DE PROTEÇÃO DE DADOS DO ARTIGO 29. **Parecer 4/2007 sobre o conceito de dados pessoais**. Bruxelas: [s. n.], 2007. p. 13.

[13] Segundo os números da época, seriam 216 milhões de indivíduos de uma população total de 248 milhões.

[14] SWEENEY, Latanya. **Simple Demographics Often Identify People Uniquely**: Data Privacy Working Paper. Pittsburgh: [s.n.], 2000. Disponível em: https://dataprivacylab.org/projects/identifiability/paper1.pdf. Acesso em: 01 mai. 2023.

a **desidentificação** de identificadores diretos[15]. A LGPD indica a diferenciação entre os conceitos de anonimização e pseudonimização ao dispor sobre o tema no contexto do regramento sobre tratamento de dados pessoais em pesquisa na área de saúde pública[16]. Conforme o texto legal do art. 13, § 4º, "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". O conceito pode ser entendido, então, como "substituição de identidade" (*ID replacement*)[17], de maneira que os identificadores diretos de certo conjunto de dados são substituídos por pseudônimos (códigos ou números randomizados, por exemplo). Assim, a identificação dos titulares dos dados permanece possível a partir do acesso ao **segredo de pseudonimização**, mantido separadamente, e adotadas as medidas de segurança e administrativas apropriadas[18].

4. ANÁLISE À LUZ DOS CONCEITOS E NORMAS FUNDAMENTAIS DE PROTEÇÃO DE DADOS PESSOAIS

4.1. Anonimização de dados e princípios de proteção de dados pessoais

Uma importante premissa adotada pelo regime de proteção de dados pessoais brasileiro é a de que, em sendo a anonimização de dados processo de remoção de identificadores diretos e indiretos de dados pessoais, os dados relativos à pessoa natural submetidos ao processo de anonimização devem ser, na origem, objeto de legítimo tratamento pelo agente responsável.[19] A afirmação possui alguns importantes

---

[15] A *desidentificação* pode ser entendida como a remoção de identificadores que diretamente identificam um indivíduo (PERSONAL DATA PROTECTION COMMISSION. **Guide to basic anonymisation**. Singapore: SG Digital-PDPC, 2022. p. 7). A noção, contudo, pode assumir sentido mais amplo na literatura, como um gênero que compreende todo processo que visa remover a associação entre um conjunto de identificadores e os titulares de dados. Cf. GARFINKEL, Simson L. **De-Identification of Personal Information**. [S.l.]: National Institute of Standards and Technology, 2015.

[16] Art. 13, *caput* e § 4º, da LGPD.

[17] Classificada como "pseudonimização tradicional" por Michèle Finck e Frank Pallas. 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. 22.

[18] EUROPEAN AGENCY FOR CYBERSECUIRTY. **Pseudonymisation techniques and best practices***:* Recommendations on shaping technology according to data protection and privacy provisions*.* [S.l.]: ENISA, 2019. p. 25. Disponível em: https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices. Acesso em: 06 mai. 2023.

[19] Nessa direção: GRUPO DE TRABALHO DE PROTEÇÃO DE DADOS DO ARTIGO 29. **Parecer 05/2014 sobre técnicas de anonimização** Bruxelas: [s. n.], 2014. p. 09.

desdobramentos. Primeiramente, fica evidenciado que o ato inicial do processo de anonimização configura operação de tratamento de dado pessoal, atraindo, assim, a aplicação de princípios e regras da LGPD.

O segundo desdobramento relevante é o de que a anonimização não é capaz de *per se* legitimar atividade de tratamento originalmente ilícita por carência de hipótese normativa que lhe dê fundamento[20]. Em outras palavras, se todo tratamento de dado pessoal deve ser legitimado por algum suporte normativo estabelecido previamente, como os previstos nos artigos 7º e 11 da LGPD, a anonimização de dados pressupõe tratamento lícito, pois não é processo apto a transformar em legítima a irregular atividade de tratamento de dados sem base legal. Se, por exemplo, num contexto de emergência sanitária um controlador que fornece aplicativo de edição de imagem e texto começa a coletar dados de geolocalização dos dispositivos de seus usuários sem qualquer hipótese legal a legitimar sua atividade, não será eventual anonimização de dados que removerá o ilícito tratamento; tais dados deverão ser eliminados.[21]

Sendo assim, por se tratar o ato inicial da anonimização de operação de tratamento de dado pessoal, deve-se levar em consideração os princípios e regras de proteção de dados pessoais aplicáveis, em especial os **princípios da finalidade**, **da adequação** e **da necessidade**. O princípio da finalidade estabelece que o tratamento de dados pessoais deverá ser realizado em consonância com propósitos legítimos, explícitos, específicos e informados ao titular quando da operação de tratamento de dados pessoais.[22] É o que prescreve o art. 6º, I, da LGPD. Isso significa dizer que, para a realização da anonimização de acordo com o regime geral de proteção de dados, deve o controlador informar com clareza que uma das finalidades da coleta dos dados pessoais é a futura

[20] Notadamente, os arts. 7 e 11 da LGPD. Em direção diversa: NEGRI, Sergio M. de C. A.; GIOVANINI, Carolina F. R. Dados não pessoais: a retórica da anonimização no enfrentamento à COVID-19 e o *privacywashing*. **Internet & Sociedade**, v. 1, n. 2, p. 126–149, 2020, p. 133.

[21] LGPD, arts. 18, IV, e 52, VI.

[22] Tal como já apontado pela ANPD anteriormente, a finalidade deve ser: "(i) **legítima**, isto é, lícita e compatível com o ordenamento jurídico, além de amparada em uma base legal, que autorize o tratamento; (ii) **específica**, de maneira que a partir da finalidade seja possível delimitar o escopo do tratamento e estabelecer as garantias necessárias para a proteção dos dados pessoais; (iii) **explícita**, isto é, expressa de uma maneira clara e precisa; e (iv) **informada**, isto é, disponibilizada em linguagem simples e de fácil compreensão e acesso ao titular dos dados". AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia orientativo –** Tratamento de dados pessoais pelo Poder Público. [S.l.]: ANPD, 2022. p. 13. Disponível em: https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/guia-poder-publico-anpd-versao-final.pdf. Acesso em: 17 mar. 2023.

anonimização[23]. Entretanto, se a finalidade de anonimização não houver sido informada originalmente, a sua realização importará "tratamento posterior"[24] ou uso secundário, que, necessariamente, deverá ser **compatível** com a finalidade inicialmente informada aos titulares dos dados[25].

Nessa linha, deve a anonimização, como tratamento posterior, observar o princípio da adequação[26], que, por sua vez, determina que a licitude da operação de tratamento depende da sua compatibilidade com a(s) finalidade(s) legítima, específica e explicitamente informada(s) ao titular dos dados, levando-se em consideração o contexto em que se realiza o tratamento. De maneira semelhante ao que já foi objeto de recomendação no "**Guia Orientativo** – Tratamento de dados pessoais pelo Poder Público", a avaliação da compatibilidade da anonimização de dados com a(s) finalidade(s) originárias deve ter em consideração, por exemplo, (i) o contexto da atividade de tratamento de dado pessoais, riscos envolvidos e outras circunstâncias relevantes do caso concreto; (ii) a existência de conexão fática ou jurídica entre a finalidade original e os objetivos do processo de anonimização; e (iii) as expectativas legítimas dos titulares e os possíveis impactos do tratamento posterior sobre seus direitos.

O princípio da necessidade é outra norma de alta relevância para a anonimização de dados. De acordo com o art. 6º, III, o tratamento de dados pessoais deverá ser limitado ao mínimo necessário para a realização de suas finalidades, abrangendo apenas os "dados pertinentes, proporcionais e não excessivos em relação às finalidades" especificadas. A necessidade do tratamento da informação exige uma avaliação preliminar direcionada a verificar se o propósito especificado pode ser alcançado com o uso mínimo de dados

---

[23] Na mesma direção: DATA PROTECTION COMMISSION. **Guidance on Anonymisation and Pseudonymisation.** [S.l.]: DPC, 2019. p. 13.

[24] De acordo com o art. 6º, I, da LGPD: "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 [...]" (grifou-se).

[25] Cf. GRUPO DE TRABALHO DE PROTEÇÃO DE DADOS DO ARTIGO 29. **Parecer 05/2014 sobre técnicas de anonimização** Bruxelas: [s. n.], 2014. p. 7-8; DATA PROTECTION COMMISSION. **Guidance on Anonymisation and Pseudonymisation.** [S.l.]: DPC, 2019. p. 13.

[26] LGPD, art. 6º, II. Na tradição do direito de proteção de dados da União Europeia (UE), as noções de "adequação" e "uso compatível" são compreendidas como elementos estruturantes do princípio da finalidade ou da limitação dos propósitos (*purpose limitation principle*) de diversas normativas. Cf. ARTICLE 29 DATA PROTECTION WORKING PARTY. **Opinion 3/2013 on purpose limitation**. Bruxelas: [s. n.], 2013. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf. Acesso em 17 mar. 2023.

pessoais ou com métodos idôneos a reduzir ou eliminar seus identificadores[27]. Dessa forma, uma vez exaurida a finalidade para a qual certos dados pessoais foram coletados, a retenção dos dados para exclusivo uso do controlador será possível desde que, à luz do princípio da necessidade, os dados sejam anonimizados[28].

Ainda neste sentido, importa chamar atenção ao fato de que a ANPD já se manifestou expressamente a respeito, porquanto "a anonimização não é uma medida de segurança impositiva, que deve ser adotada em todo e qualquer tratamento de dados pessoais". A pertinência da adoção do processo de anonimização decorre de um juízo de necessidade à luz da(s) finalidade(s) especificada(s) para o tratamento de dados na situação concreta.[29]

4.2. Os conceitos de dado pessoal e dado anonimizado

Para compreender o processo de anonimização e os dados anonimizados, que com esse processo se busca produzir, é imprescindível conhecer o conceito de dado pessoal e seus contornos. Aliás, entender este conceito, por consequência, implica estabelecer os limites do que não é informação pessoal, ou seja, a noção de dado não pessoal. Esta pode ser entendida como gênero que abrange duas espécies, os **dados anônimos** e os **dados anonimizados**.[30]

---

[27] D'ORAZIO, Roberto. Il principio di necessità nel trattamento dei dati personali. *In*: CUFFARO, Vincenzo; D'ORAZIO, Roberto; RICCIUTO, Vincenzo (Org.). **Il Codice del trattamento dei dati personali**. Torino: Giappichelli, 2007. p. 21. Esse era um preceito garantido no art. 3º do reformado "Código em matéria de proteção de dados pessoais" da Itália, com muita aproximação à norma brasileira vigente. Na mesma direção: AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. **10 malentendidos relacionados com la anonimización.** Disponível em: https://www.aepd.es/documento/10-malentendidos-anonimizacion.pdf. Acesso em 07 ago. 2023.

[28] LGPD, art. 16, IV

[29] AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Nota Técnica nº 46/2022/CGF/ANPD.** Disponível em: https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/sei_00261-000730_2022_53-nt-46.pdf. Acesso em: 07 ago. 2023; AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia Orientativo – Tratamento de dados pessoais para fins acadêmicos e para a realização de estudos e pesquisas**. Brasília: ANPD, 2023. p. 41.

[30] Numa direção similar, Kerstin N. Vokinger, Daniel J. Stekhoven e Michael Krauthammer diferenciam entre **dados anonimizados** (reversíveis e irreversíveis) e **dados anônimos**, sendo esta última categoria definida como dados que foram coletados numa forma anônima ou foram agregados e a reidentificação de indivíduos não é possível (VOKINGER, Kerstin N.; STEKHOVEN, Daniel J.; KRAUTHAMMER, Michael. Lost in Anonymization — A Data Anonymization Reference Classification Merging Legal and Technical Considerations. **Journal of Law, Medicine and Ethics**, v. 48, n. 1, p. 228–231, 2020, p. 229). Documento da AEPD

Dados anônimos são dados que, desde sua origem, não se relacionam a pessoa natural identificada ou identificável. É o que se dá, por exemplo, com os dados sobre condições meteorológicas processados e divulgados em aplicação de internet. **Dados anonimizados** são, por sua vez, aqueles dados inicialmente vinculados à pessoa natural, mas que foram posteriormente submetidos a processo de anonimização a partir de técnicas ou paradigmas como generalização e privacidade diferencial. Em razão da remoção dos identificadores diretos e indiretos, os dados perdem, a princípio, o caráter pessoal. São justamente os identificadores os pontos que estabelecem o vínculo de identificabilidade do dado com uma pessoa natural. A identificabilidade é, então, um elemento central do conceito de dado pessoal.

Entre as diferentes perspectivas de política regulatória em proteção de dados[31], o direito brasileiro adotou com a LGPD a perspectiva expansionista[32], a qual, por sua vez, inspirou a positivação do **conceito amplo** no art. 5º, I, da referida lei[33]. Isso significa que o conceito jurídico de dado pessoal estende seu alcance para além da pessoa natural identificada, ou precisamente individualizada a partir de identificadores diretos e indiretos. É igualmente considerada informação pessoal aquela relativa a pessoa natural **identificável**.

Na concepção ampla de dados pessoais – aderida por diversas jurisdições e órgãos internacionais, a exemplo da União Europeia (UE), Conselho da Europa, Argentina e Colômbia – abordagens diferentes podem ser escolhidas a depender da importância que se atribui à atuação e aos esforços de terceiros na avaliação da

---

[31] 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. 1814–189, dez. 2011. De acordo com estes autores estadunidenses, para a abordagem expansionista é irrelevante se os dados já foram vinculados a uma determinada pessoa natural ou se poderão ser vinculadas no futuro; é uma visão que trata os dados identificados e identificáveis como equivalentes.

[32] MACHADO, Diego; DONEDA, Danilo. Proteção de dados pessoais e criptografia: tecnologias criptográficas entre anonimização e pseudonimização de dados *In*: DONEDA, Danilo (coord.). **A regulação da criptografia no direito brasileiro**. São Paulo: Revista dos Tribunais, 2018. p. 104; BIONI, Bruno R. **Proteção de dados pessoais:** a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019. p. 74-75

[33] Cf. MARTINS, Guilherme M.; LONGHI, João Victor R.; FALEIROS JÚNIOR, José Luiz de M. **Comentários à Lei Geral de Proteção de Dados Pessoais**: Lei 13.709/2018. Induiutaba: Editora Foco, 2022. p. 153. Cumpre ressaltar, contudo, que no julgamento do Recurso Extraordinário 673.707/MG, pelo Supremo Tribunal Federal, o Ministro Luiz Fux já então sinalizara em seu voto pela adoção, no campo da aplicação do *habeas data*, de uma noção ampla de dados (pessoais): "Encarta-se, assim, no conceito mais amplo de arquivos, bancos ou registro de dados, que devem ser entendidos em seu sentido mais lato, abrangendo tudo que diga respeito ao interessado, seja de modo direto ou indireto". BRASIL. Supremo Tribunal Federal (Tribunal Pleno). Recurso Extraordinário n. 673.707/MG. Relator: Min. Luiz Fux. Brasília, 07/02/2017. **Diário de Justiça eletrônico**, Brasília, DF, 30/09/2015.

identificabilidade da pessoa natural[34]. Denomina-se abordagem **relativa** ou **subjetiva** a orientação que tem em vista apenas as habilidades, os esforços individuais e uso de meios próprios do agente que trata os dados em questão (*v.g.*, o controlador que anonimizou certa base de dados), ou seja, é abordagem que, na apreciação da identificabilidade, desconsidera a atividade e esforços de outrem[35]. De outra parte, a perspectiva que considera o potencial de objetiva identificação não só por esforços e uso de meios próprios do responsável pela operação de tratamento, mas também de qualquer outra pessoa ou ente (*v.g.*, terceiros com quem se compartilhou base de dados anonimizada), é chamada de **absoluta** ou **objetiva**.

Na discussão a respeito dos tipos de abordagem dentro do conceito amplo, há pertinência de uma breve análise sobre como o direito da UE trata a questão, haja vista a tradição jurídica conquistada na matéria e, especialmente para fins comparativos segundo o método funcionalista, a aproximação com o direito brasileiro. A comparação será especialmente relevante para a compreensão das noções de **meios próprios** e **esforços razoáveis** no item 4.4.

---

[34] Cf. BORGESIUS, Frederik Zuiderveen. Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. **Computer Law and Security Review**, v. 32, n. 2, p. 256–271, 2016, p. 8-9; 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, p. 163-177, 2016, p. 165; 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. 40–81, 2018, p. 46, 64; 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. 17; PAHLEN-BRANDT, Ingrid. Datenschutz braucht scharfe Instrumente Beitrag zur Diskussion um „personenbezogene Daten“. **Datenschutz und Datensicherheit – DuD,** v. 32, p. 34–40, jan. 2008, p. 38.

[35] Esta parece ter sido a abordagem adotada pelo Comitê de Ministros do Conselho da Europa na Recomendação CM/Rec(2021)8, cujo art. 1. 1, *a*, conceitua dado pessoal da seguinte forma: ““Personal data” means any information relating to an identified or identifiable natural person (“data subject”). An individual is not considered “identifiable” if identification requires unreasonable time, *resources or effort in relation to the means at the disposal of the controller*” (grifou-se). COUNCIL OF EUROPE. **Recommendation CM/Rec (2021)8 of the Committee of Ministers to member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling.** Disponível em: https://search.coe.int/cm/pages/result_details.aspx?objectid=0900001680a46147. Acesso em: 12 abr. 2023.

O direito de proteção de dados da UE, seja com base no texto da revogada Diretiva n. 95/46/CE[36] ou do vigente Regulamento n. 2016/679[37], possui convergências com a teoria absoluta ou objetiva, visto que leva em consideração para a identificabilidade do titular dos dados os meios e esforços empregados não apenas pelo controlador, mas como por qualquer terceiro[38]. No considerando n. 26 da revogada diretiva europeia e do vigente Regulamento Geral de Proteção de Dados da UE, são tratados como meios suscetíveis de serem utilizados para (re)identificação de titular de dados aqueles adotados "**seja pelo responsável pelo tratamento, seja por qualquer outra pessoa**". A título de exemplo, pode ser citado o processo C-582/14, *Patrick Breyer v. Bundesrepublik Deutschland*, julgado pelo Tribunal de Justiça da União Europeia (TJUE) no ano de 2016[39]. No caso, analisou-se se o endereço IP deve ser compreendido como dado pessoal à luz do conceito amplo adotado pela Diretiva n. 95/46/CE, bem como pelo Regulamento n. 2016/679, num cenário em que o provedor de aplicação de internet realiza operações de tratamento com o endereço IP, mas é o provedor de conexão à internet quem possui os dados cadastrais aptos a efetivamente identificar o usuário. Daí o questionamento: o endereço IP é dado pessoal? Ou será que poderia ser tratado como dado anônimo?

O TJUE afirmou nos fundamentos da decisão judicial que, porquanto o Considerando n. 26 da Diretiva n. 95/46/CE faz referência aos meios suscetíveis de ser utilizados quer pelo responsável pelo tratamento quer por qualquer outra pessoa, entende-se que sua redação sugere não ser imprescindível que todas as informações que possibilitam identificar o titular de dados estejam nas mãos de um único agente. Sendo assim, para a corte europeia, o fato de os dados cadastrais (e registros de acesso à internet)

---

[36] UNIÃO EUROPEIA. Directiva 95/46/CE do Parlamento Europeu e do Conselho, de 24 de outubro de 1995, relativa à protecção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A31995L0046. Acesso em: 18 abr. 2023.

[37] UNIÃO EUROPEIA. Regulamento (UE) nº 2016/679 do Parlamento Europeu e do Conselho, de 23 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). **Jornal Oficial da União Europeia**, Estrasburgo, 04/05/2016. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679. Acesso em: 18 abr. 2023.

[38] Começam a surgir, no entanto, abordagens téorico-regulatórias intermediárias, a exemplo da proposta pelo *Information Commissioner's Office,* do Reino Unido, que leva em consideração não qualquer pessoa, mas o "intruso motivado" com "razoável competência" para lançar mão de conhecimento, meios e recursos para identificar alguém. INFORMATION COMMISSIONER`S OFFICE. **Anonymisation:** managing data protection risk code of practice. Wilmslow: ICO, 2012. p. 22-23.

[39] UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia. Grande Secção. **Processo C-582/14, Patrick Breyer v. Bundesrepublik Deutschland**. Luxemburgo, 19 out. 2016. Disponível em: https://curia.europa.eu/juris/document/document.jsf?docid=184668&doclang=PT. Acesso em: 03 mar. 2023.

necessários para identificar o usuário de um sítio eletrônico não ser objeto de tratamento pelo provedor de aplicação, mas sim pelo provedor de conexão à internet contratado pelo usuário, não é capaz de excluir o caráter pessoal dos endereços IP tratados pelo provedor de aplicação de internet.

Aliás, o entendimento fixado pelo tribunal europeu é compatível com o sistema brasileiro e com a estrutura normativa da Lei n. 12.695/2014, o Marco Civil da Internet, que prevê deveres de guarda de registros (ou *logs*) de conexão e de acesso à aplicação, respectivamente, a provedores de conexão e a provedores de aplicação de internet. Nessa linha, o Superior Tribunal Justiça afirmou que, "[...] tem-se, na prática, uma repartição das informações de navegação: i) o provedor de conexão, ao habilitar um terminal para envio e recebimento de dados, atribui a ele um IP e registra o momento em que iniciada, interrompida e encerrada a conexão, e ii) cada provedor de aplicação registra o acesso dos IPs, momento de início e final, à sua própria aplicação. Desse modo, a totalidade da navegação de cada internauta dependerá da remontagem de cada uma das aplicações acessadas ao longo de uma única conexão"[40].

Muito embora inicialmente prevaleça a abordagem objetiva, estipulou-se importante critério para limitar a abrangência do conceito de dado pessoal: o critério dos **meios suscetíveis de ser razoavelmente utilizados**. Na ausência desse limitador, caso uma única pessoa ou entidade, independentemente dos meios usados, seja capaz de identificar alguém a partir de certo dado, esse dado seria qualificado como de caráter pessoal pela abordagem objetiva ou absoluta. Nesse sentido, o considerando n. 26 da Diretiva n. 95/46/CE, estabelecia que "*para determinar se uma pessoa é identificável, importa considerar o conjunto dos meios susceptíveis de serem razoavelmente utilizados*".

No atual Regulamento n. 2016/679 da UE, o tema é abordado por considerando de mesmo número, que, além de fazer menção aos "meios suscetíveis de ser razoavelmente utilizados", indica aos intérpretes da legislação europeia a relevância de aspectos objetivos a serem avaliados:

> Para determinar se uma pessoa singular é identificável, importa considerar todos os meios suscetíveis de ser razoavelmente utilizados, tais como a seleção, quer pelo responsável pelo tratamento quer por outra pessoa, para identificar direta ou indiretamente a pessoa singular. Para determinar se há uma

[40] BRASIL. Superior Tribunal de Justiça (Terceira Turma). Recurso Especial n. 1.784.156/SP. Relatora: Min. Aurélio Bellizze. Brasília, 05/11/2019. **Diário de Justiça eletrônico**, Brasília, DF, 21/11/2019.

> probabilidade razoável de os meios serem utilizados para identificar a pessoa singular, importa considerar todos os fatores objetivos, como os custos e o tempo necessário para a identificação, tendo em conta a tecnologia disponível à data do tratamento dos dados e a evolução tecnológica[41].

A inclusão de uma lista exemplificativa de aspectos objetivos a serem considerados[42], tais como o tempo necessário para a (re)identificação de titular de dados em relação a certa base de dados anonimizada com informações sobre o deslocamento de usuários de transporte público de uma metrópole numa série temporal, foi decerto impulsionada por documentos como o Parecer n. 04/2007 do extinto Grupo de Trabalho do Artigo 29. Neste parecer, o antigo órgão consultivo fez referência ao considerando n. 26 da Diretiva n. 95/46/CE, a fim de esclarecer que este critério não diz respeito a alguma possibilidade hipotética de (re)identificação de alguém, mas é **dinâmico** e **contextual**: todos os meios disponíveis para (re)identificação devem ser levados em conta no momento da análise, haja vista os objetivos da atividade de tratamento, tempo e custos da identificação, estado da arte e desenvolvimento previsíveis da tecnologia e técnicas de reidentificação durante o período de tratamento dos dados etc.[43]

A jurisprudência do TJUE criou outro fator objetivo a ser levado em consideração ao analisar os meios que podem razoavelmente ser utilizados: a licitude dos meios. Desse modo, meios ilegais ou que implicam infração ao direito vigente aplicável configuram meios e esforços não razoáveis de ser empregados. Tal posição foi adotada no caso C-582/14, *Patrick Breyer v. Bundesrepublik Deutschland*, e recentemente confirmada pela corte no Processo T-557/20, *Conselho Único de Resolução v. Autoridade Europeia de Proteção de Dados*.[44]

É importante destacar que o referido critério dos meios passíveis de ser razoavelmente utilizados para a identificação do titular dos dados tem sido interpretado

---

[41] UNIÃO EUROPEIA. Regulamento (UE) nº 2016/679 do Parlamento Europeu e do Conselho, de 23 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). **Jornal Oficial da União Europeia**, Estrasburgo, 04/05/2016. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679. Acesso em: 18 abr. 2023.

[42] KUNER, Christopher; BYGRAVE, Lee A.; DOCKSEY, Christopher. **The EU General Data Protection Regulation:** A Commentary. 1. ed. Oxford: Oxford University Press, 2020. P. 108.

[43] GRUPO DE TRABALHO DE PROTEÇÃO DE DADOS DO ARTIGO 29. **Parecer 4/2007 sobre o conceito de dados pessoais**. Bruxelas: [s. n.], 2007. p. 16.

[44] UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia. Oitava Secção alargada. **Processo T-557/20, Conselho Único de Resolução v. Autoridade Europeia de Proteção de Dados**. Luxemburgo, 26 abr. 2023. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:62020TJ0557. Acesso em: 09 mai. 2023.

por diversos autores como elemento normativo que introduz uma análise de risco de identificabilidade[45], ou um teste de risco de identificabilidade[46]. Seria, assim, a aplicação de um **modelo baseado em riscos** à delimitação do que configura dado pessoal (ou dado não pessoal) e, portanto, ao âmbito material de incidência de regimes gerais de proteção de dados pessoais como a LGPD.

Por coerência lógica e sistemática, a compreensão de dado anonimizado, ou seja, "dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento" (LGPD, art. 5º, III), necessariamente é moldada pela conceituação jurídica de dado pessoal e pela abordagem adotada. A LGPD, muito embora deixe clara a opção pelo **conceito amplo**, no art. 5º, I, não fornece parâmetros sobre a abordagem ou teoria abarcada, se relativa ou objetiva – deixando espaço, aliás, para a articulação de abordagens intermediárias.[47] É o conceito de dado anonimizado[48] e a previsão sobre a reversibilidade do processo de anonimização que apresentam os parâmetros dos meios e esforços razoáveis, inclusive daqueles que não sejam exclusivamente próprios[49] (vide item 4.4).

4.3. O processo de anonimização de dados e riscos de (re)identificação

O processo de anonimização aplicado à base de dados pessoais visa a atuar sobre os identificadores diretos e indiretos a fim de obter, ao longo de sua contínua duração, dados que, irreversivelmente, não identificam nem possam identificar titulares de dados – dados anonimizados, portanto. Trata-se de processo que se desenvolve por meio da utilização de técnicas diversificadas – a exemplo da supressão, generalização, permutação e perturbação – cuja pertinência é justificada de acordo com as características e outros aspectos contextuais da base de dados que o agente de tratamento pretende anonimizar.

---

[45] 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. 1814–189, dez. 2011, p. 1878.
[46] 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. 13-14.
[47] Vide nota de rodapé nº 37.
[48] LGPD, art. 5º, III.
[49] LGPD, art. 12.

Isso porque além de a LGPD não impor qualquer tipo de técnica de anonimização específica, não há qualquer metodologia universalmente aplicável[50].

De acordo com o atual estado da arte e desenvolvimento científico, pode-se afirmar a existência de um consenso científico sobre a impraticabilidade de um cenário de ausência de risco de reidentificação[51] nas situações de tratamento de dados anonimizados[52]. Tendo em vista o enorme volume de dados auxiliares disponibilizados publicamente via internet e o desenvolvimento da capacidade de processamento e análise de algoritmos de reidentificação, é fundada a afirmativa de que sempre haverá fatores de risco de reidentificação. Nesse sentido, a adoção de **modelo baseado em riscos** relacionado à identificabilidade de dados também se mostra pertinente na avaliação da robustez do processo de anonimização, que não pode ser episódica, mas sim iterativa, porquanto os novos riscos podem advir ao longo do tempo na medida dos avanços tecnológicos e da quantidade de dados auxiliares, por exemplo.

Os riscos de reidentificação de dados anonimizados são expressos, em linguagem técnica, como possíveis **ataques de reidentificação**. O termo "ataque" é tomado por empréstimo da literatura especializada em segurança computacional, em que a avaliação do nível de segurança de determinado sistema computacional ou algoritmo de cifragem ocorre a partir do uso da figura de um hipotético "atacante" que possui certas habilidades, conhecimento ou acesso[53]. "Uma avaliação de risco envolve a catalogação da variedade de potenciais atacantes, e, para cada um, a probabilidade de sucesso"[54]. Cumpre ressaltar

---

[50] Cf. NARAYANAN, Arvind; SHMATIKOV, Vitaly. Myths and fallacies of "personally identifiable information". **Communications of the ACM**, v. 53, n. 6, p. 24-26, jun. 2010, p. 26.

[51] Denonima-se "risco de reidentificação" o risco de identificação incidente sobre dados anonimizados, isto é, aqueles dados resultantes da implementação de processo de anonimização.

[52] No arco de mais duas décadas, vários estudos científicos demonstram os vários riscos de reidentificação no tratamento de bases de dados assumidas a princípio como anonimizadas. Dentre alguns desses trabalhos pode-se citar: SWEENEY, Latanya. Simple Demographics Often Identify People Uniquely. **Carnegie Mellon University**, Data Privacy Working Paper 3, Pittsburgh, 2000; NARAYANAN, Arvind; SHMATIKOV, Vitaly. **How to break anonymity of the Netflix Prize dataset**. 2007. Disponível em: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.3581&rep=rep1&type=pdf. Acesso em: 10 mai. 2023; WONDRACEK, Gilbert et al. A practical attack to de-anonymize social network users. **Proceedings – IEEE Symposium on Security and Privacy**, p. 223–238, 2010; DE MONTJOYE, Yves A. et al. Unique in the Crowd: The privacy bounds of human mobility. **Scientific Reports**, v. 3, p. 1–5, 2013; ROCHER, L.; HENDRICKX, J. M.; DE MONTJOYE, Y. A. Estimating the success of re-identifications in incomplete datasets using generative models. **Nature Communications**, v. 10, n. 1, 2019.

[53] A figura do "atacante" muito se aproxima do "intruso" (*intruder*) a que a autoridade de proteção de dados da Irlanda se refere: DATA PROTECTION COMMISSION. **Guidance on Anonymisation and Pseudonymisation.** [S.l.]: DPC, 2019. p. 8-10.

[54] GARFINKEL, Simson L. **De-Identification of Personal Information**. [S.l.]: National Institute of Standards and Technology, 2015. p. 9.

que essa noção de "atacante" não se confunde com atores que praticam crimes ou atos antijurídicos. Basta considerar o exemplo de pesquisadores que avaliam a robustez de base de dados anonimizada compartilhada publicamente frente a certos algoritmos de reidentificação com o uso de dados auxiliares disponibilizados em bases de acesso público[55].

Alguns exemplos de ataques ou riscos de reidentificação que podem ser mencionados são (i) a distinção (*singling out*), (ii) a possibilidade de ligação (*linkability*), e (iii) a inferência[56]. A **distinção** consiste na possibilidade de se isolar alguns ou todos os registros que destacam um indivíduo numa base de dados. A **possibilidade de ligação** é definida pela capacidade de se estabelecer uma conexão entre pelo menos dois registros relativos ao mesmo indivíduo ou mesmo grupo de pessoas. Já o risco de **inferência** diz respeito à possibilidade de inferir, com uma significativa probabilidade, o valor de um atributo a partir dos valores de um conjunto de outros atributos.[57]

4.4. As noções de *esforços razoáveis* e *meios próprios*

A compreensão do processo de anonimização e os critérios a serem considerados para avaliar sua reversibilidade a partir dos riscos de reidentificação, requer,

---

[55] Basicamente, foi o que os pesquisadores Arvind Narayanan e Vitaly Shmatikov fizeram em estudo sobre a base de dados do *"Netflix Prize"*: NARAYANAN, Arvind; SHMATIKOV, Vitaly. **How to break anonymity of the Netflix Prize dataset**. 2007. Disponível em: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.3581&rep=rep1&type=pdf. Acesso em: 10 mai. 2023.

[56] CAVOUKIAN, Ann; EMAM, Khaled El. Dispelling the Myths Surrounding Anonymization Remains a Strong Tool for Protecting Privacy. **Information and Privacy Commissioner,** Ontario, Canada, n. June, 2011. Disponível em: https://www.ipc.on.ca/wp-content/uploads/2016/11/anonymization.pdf. Acesso em: 06 abr. 2023.

[57] Para melhor explicar o risco de inferência, considere o uso de dados agregados de geolocalização. A reidentificação dos usuários de dispositivos móveis é um risco que, não obstante minorado pela agregação de dados de localização, não é eliminado. Os cientistas da computação Pyergelis, Troncoso e De Cristofaro apontam que a elaboração de modelos de mobilidade com esse tipo de dado coletado por certo período de tempo é sujeito a ameaça capaz de, *por inferência*, reconhecer a contribuição de uma pessoa na formação de um agregado de geolocalização, possuindo o adversário dados auxiliares. O ataque, denominado *membership inference attack* (MIA), possui significativas implicações, conforme destacam os autores: (i) o só fato de se concluir que os dados de alguém faz parte de um agregado pode constituir informação sensível; e (ii) esse tipo de ataque é um primeiro passo para ulteriores inferências que visam obter informações adicionais sobre indivíduos, tais como seu perfil de mobilidade e/ou suas trajetórias a partir dos dados agregados. PYRGELIS, Apostolos; TRONCOSO, Carmela; DE CRISTOFARO, Emiliano. Measuring Membership Privacy on Aggregate Location Time-Series. **Proceedings of the ACM on Measurement and Analysis of Computing Systems**, v. 2, n. 4, 2020, p. 1.

necessariamente, a interpretação de dois termos previstos textualmente no artigo 12, *caput,* da LGPD[58], a saber, **"esforços razoáveis"** e **"meios próprios"**.

O primeiro configura um **conceito jurídico indeterminado normativo***,* é dizer, um conceito em larga medida incerto em seu conteúdo e extensão, dependente de preenchimento valorativo pelo aplicador do Direito[59]. Isso significa em termos práticos que a ANPD, como intérprete e aplicadora da LGPD, deve preencher, com elementos e critérios pertinentes com o caso concreto, a noção de "esforços razoáveis", dentro do sentido literal possível e em coesão com o contexto significativo da lei, que, aliás, prevê no § 1º do art. 12[60], relevantes parâmetros interpretativos.

Além da referência à própria LGPD, se mostra útil a atenção com certa tradição jurídica da proteção de dados pessoais, visto que o sistema nacional buscou inspiração em outras leis e experiências jurídicas. Nesse sentido, pode-se estabelecer uma direta conexão entre a noção de "esforços razoáveis" com o critério expressamente previsto no direito da UE dos "meios suscetíveis de ser razoavelmente utilizados" (vide item 2.2.2.) para tornar uma pessoa natural reidentificável.

Observando semelhante direção da que foi adotada na legislação de proteção de dados da UE, a LGPD estabelece no artigo 12, § 1º, um **rol exemplificativo** de aspectos objetivos que devem ser sopesados pelo intérprete ao preencher (ou determinar), nas situações concretas, o conteúdo normativo do que é esforço razoável. Conforme o texto da lei, "na 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".

Na análise dos fatores **custo e tempo** necessários para a possibilidade de reidentificação dos titulares e reversão do processo de anonimização, deve se considerar, por exemplo, as medidas administrativas de controle e gestão de acesso interno na

---

[58] "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".

[59] ENGISCH, Karl. **Introdução ao pensamento jurídico**. 8. ed. Trad. João Baptista Machado. Lisboa: Fundação Calouste Gulbenkian, 2001. p. 210-213.

[60] "§ 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".

organização do agente de tratamento responsável pelo processo de anonimização, bem como as técnicas implementadas.[61]

Outro fator importante para a compreensão dos (meios e) esforços razoáveis para reidentificação ou reversibilidade do processo de anonimização é sua **licitude**. O uso desse fator objetivo implica dizer que a prática de crimes cibernéticos ou atos ilícitos configuram meios e esforços irrazoáveis para a reidentificação ou reversão do processo de anonimização. Nesse sentido, a violação de atos normativos, acordos, contratos, atos jurídicos de natureza negocial ou estatutária como políticas internas que estabelecem regras de conduta a prestadores de serviço, trabalhadores ou servidores, podem consistir em meios e esforços irrazoáveis para a reidentificação ou reversão de processo de anonimização.

Entretanto, diferentemente do que se viu a respeito da noção de "esforços razoáveis", o conceito de **meios próprios** tem conteúdo mais delimitado, não caracterizando-se, assim, como conceito juridicamente indeterminado. A partir do seu teor literal e possível pode-se afirmar que são meios próprios as habilidades, os dados, instrumentos e técnicas disponíveis ao próprio agente de tratamento responsável pela anonimização.

É importante ressaltar que, a partir da redação legal do art. 12, *caput*, da LGPD, depreende-se que avaliação da possibilidade de reidentificação de dados e reversão do processo de anonimização deve ter em consideração não apenas o uso de meios próprios do agente de tratamento responsável pela anonimização, mas também a atuação de outras pessoas ou entidades que, com meios e esforços razoáveis, podem reidentificar conjunto de dados anonimizados. Isso significa, na verdade, que a LGPD aponta para a teoria objetiva do conceito amplo de dado pessoal.

No entanto, haja vista que na avaliação da possibilidade de reidentificação são analisados os meios e esforços razoáveis do próprio agente de tratamento responsável

---

[61] Neste sentido, a ANPD já teve a oportunidade de se manifestar a respeito em Nota Técnica elaborada no caso envolvendo o tratamento de microdados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP): "5.27. A avaliação relativa à eventual reversão dos dados e aos seus impactos deve se basear em evidências e em cenários que considerem aspectos objetivos da realidade. Afastam-se, assim, análises meramente especulativas, baseadas em cenários irreais, de difícil ou improvável ocorrência ou, ainda, que desconsiderem limitações práticas, decorrentes de custos muito elevados ou de meios técnicos de disponibilidade restrita." AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Nota Técnica nº 46/2022/CGF/ANPD.** Disponível em: https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/sei_00261-000730_2022_53-nt-46.pdf. Acesso em: 07 ago. 2023.

pelo processo de anonimização e os meios e esforços razoáveis de outrem, deve ser aplicado um **modelo ou abordagem baseada em riscos de (re)identificação** ao longo do contínuo processo de anonimização[62].

CONSIDERAÇÕES FINAIS

O presente estudo buscou delinear as bases jurídico-normativas do processo de anonimização de dados na LGPD. Não se trata de um trabalho que esgota o tema da anonimização. Ao contrário, lança alicerces para futuras contribuições e orientações da ANPD a todos atores envolvidos no ecossistema brasileiro de proteção de dados pessoais.

Entre os pontos desenvolvidos neste estudo introdutório, merecem destaque:

a) a caraterização da anonimização de dados como processo sob a ótica jurídica e a técnica;

b) o ato inicial do processo de anonimização configura operação de tratamento de dado pessoal, fazendo incidirem princípios e regras do regime da LGPD, tais como os princípios da finalidade e da adequação;

c) conceitos de dado pessoal e dado anonimizado na LGPD e no direito brasileiro devem ser interpretados conjunta e sistematicamente e em conformidade com o caráter dinâmico e contextual que possuem;

d) a anonimização não é uma medida impositiva, que deve ser adotada em todo e qualquer tratamento de dados pessoais;

e) o critério dos esforços razoáveis aponta para a necessidade de análise de riscos de (re)identificabilidade; e

f) a avaliação do contínuo processo de anonimização deve se dar de acordo com um modelo baseado em riscos.

---

[62] BIONI, Bruno. Compreendendo o conceito de anonimização e dado anonimizado. **Cadernos Jurídicos**, ano 21, n. 53, p. 191-201, jan./mar. 2020, p. 197; RUBINSTEIN, Ira S.; HARTZOG, Woodrow. Anonymization and risk. **Washington Law Review**, v. 91, n. 2, p. 703–760, 2016, p. 747 et seq.

REFERÊNCIAS

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. **Orientaciones y garantias em los procedimentos de anonimización de datos personales**. [S.l.]: AEPD, 2016. Disponível em: https://www.aepd.es/es/documento/guia-orientaciones-procedimientos-anonimizacion.pdf. Acesso em: 02 mar. 2023.

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. **10 malentendidos relacionados com la anonimización.** Disponível em: https://www.aepd.es/documento/10-malentendidos-anonimizacion.pdf. Acesso em 07 ago. 2023.

AGENCIA DE GOBIERNO ELECTRÓNICO Y SOCIEDAD DE LA INFORMACIÓN Y DEL CONOCIMIENTO. **Guía sobre Anonimización de Datos**. [S.l]: AGESIC, 2020. Disponível em: https://www.gub.uy/unidad-reguladora-control-datos-personales/comunicacion/publicaciones/guia-criterios-disociacion-datos-personales/guia-criterios-disociacion. Acesso em: 01 mai. 2023.

ARTICLE 29 DATA PROTECTION WORKING PARTY. **Opinion 3/2013 on purpose limitation**. Bruxelas: [s. n.], 2013. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf. Acesso em 17 mar. 2023.

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia orientativo –** Tratamento de dados pessoais pelo Poder Público. [S.l.]: ANPD, 2022. Disponível em: https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/guia-poder-publico-anpd-versao-final.pdf. Acesso em: 17 mar. 2023.

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Nota Técnica nº 46/2022/CGF/ANPD.** Disponível em: https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/sei_00261-000730_2022_53-nt-46.pdf. Acesso em: 07 ago. 2023.

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia Orientativo – Tratamento de dados pessoais para fins acadêmicos e para a realização de estudos e pesquisas**. Brasília: ANPD, 2023. Disponível em: https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/documentos-de-publicacoes/web-guia-anpd-tratamento-de-dados-para-fins-academicos.pdf. Acesso em: 07 ago. 2023.

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

BIONI, Bruno. Compreendendo o conceito de anonimização e dado anonimizado. **Cadernos Jurídicos**, ano 21, n. 53, p. 191-201, jan./mar. 2020.

BORGESIUS, Frederik Zuiderveen. Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. **Computer Law and Security Review**, v. 32, n. 2, p. 256–271, 2016.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 09 mai. 2023.

BRASIL. **Lei nº 14.534, de 11 de janeiro de 2023.** Altera as Leis nºs 7.116, de 29 de agosto de 1983, 9.454, de 7 de abril de 1997, 13.444, de 11 de maio de 2017, e 13.460, de 26 de junho de 2017, para adotar número único para os documentos que especifica e para estabelecer o Cadastro de Pessoas Físicas (CPF) como número suficiente para identificação do cidadão nos bancos de dados de serviços públicos. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2023/lei/l14534.htm. Acesso em: 09 mai. 2023.

BRASIL. Supremo Tribunal Federal (Tribunal Pleno). Recurso Extraordinário n. 673.707/MG. Relator: Min. Luiz Fux. Brasília, 07/02/2017. **Diário de Justiça eletrônico**, Brasília, DF, 30/09/2015.

BRASIL. Superior Tribunal de Justiça (Terceira Turma). Recurso Especial n. 1.784.156/SP. Relatora: Min. Aurélio Bellizze. Brasília, 05/11/2019. **Diário de Justiça eletrônico**, Brasília, DF, 21/11/2019.

CAVOUKIAN, Ann; EMAM, Khaled El. Dispelling the Myths Surrounding Anonymization Remains a Strong Tool for Protecting Privacy. **Information and Privacy Commissioner, Ontario, Canada**, n. June, 2011. Disponível em: https://www.ipc.on.ca/wp-content/uploads/2016/11/anonymization.pdf. Acesso em: 06 abr. 2023.

COUNCIL OF EUROPE. **Recommendation CM/Rec (2021)8 of the Committee of Ministers to member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling.** Disponível em: https://search.coe.int/cm/pages/result_details.aspx?objectid=0900001680a46147. Acesso em: 12 abr. 2023.

COUTO E SILVA, Clóvis V. do. **A obrigação como processo**. Rio de Janeiro: Editora FGV, 2006.

DATA PROTECTION COMMISSION. **Guidance on Anonymisation and Pseudonymisation.** [S.l.]: DPC, 2019.

DE MONTJOYE, Yves A. et al. Unique in the Crowd: The privacy bounds of human mobility. **Scientific Reports**, v. 3, p. 1–5, 2013.

DIDIER JR., Fredie. **Curso de direito processual civil**. 17. ed. rev., atual. e amp. Salvador: JusPodivm, 2015. v. 1.

DONEDA, Danilo; MENDES, Laura Schertel. Um perfil da nova Lei Geral de Proteção de Dados brasileira. *In*: BELLI, Luca; CAVALLI, Olga (Orgs.). **Governança e Regulações da Internet na América Latina**. Rio de Janeiro: FGV Direito Rio, 2018. P. 309–324.

D'ORAZIO, Roberto. Il principio di necessità nel trattamento dei dati personali. *In*: CUFFARO, Vincenzo; D'ORAZIO, Roberto; RICCIUTO, Vincenzo (Org.). **Il Codice del trattamento dei dati personali**. Torino: Giappichelli, 2007.

ENGISCH, Karl. **Introdução ao pensamento jurídico**. 8. ed. Trad. João Baptista Machado. Lisboa: Fundação Calouste Gulbenkian, 2001.

EUROPEAN AGENCY FOR CYBERSECUIRTY. **Pseudonymisation techniques and best practices***:* Recommendations on shaping technology according to data protection and privacy provisions. [S.l.]: ENISA, 2019.

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.

GARFINKEL, Simson L. **De-Identification of Personal Information**. [S.l.]: National Institute of Standards and Technology, 2015.

GRUPO DE TRABALHO DE PROTECÇÃO DE DADOS DO ARTIGO 29. **Parecer 4/2007 sobre o conceito de dados pessoais**. Bruxelas: [s. n.], 2007. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_pt.pdf. Acesso em: 10 mar. 2023.

GRUPO DE TRABALHO DE PROTEÇÃO DE DADOS DO ARTIGO 29. **Parecer 05/2014 sobre técnicas de anonimização** Bruxelas: [s. n.], 2014.

GUSTIN, Miracy B. S.; DIAS, Maria Tereza Fonseca; NICÁCIO, Camila Silva. **(Re)pensando a pesquisa jurídica***:* teoria e prática. 5. ed. rev., amp. e atual. São Paulo: Almedina, 2020.

INFORMATION COMMISSIONER`S OFFICE. **Anonymisation:** managing data protection risk code of practice. Wilmslow: ICO, 2012. Disponível em: https://ico.org.uk/media/1061/anonymisation-code.pdf. Acesso em: 16 abr. 2023.

KUNER, Christopher; BYGRAVE, Lee A.; DOCKSEY, Christopher. **The EU General Data Protection Regulation:** A Commentary. 1. ed. Oxford: Oxford University Press, 2020.

MACHADO, Diego; DONEDA, Danilo. Proteção de dados pessoais e criptografia: tecnologias criptográficas entre anonimização e pseudonimização de dados *In*: DONEDA, Danilo (coord.). **A regulação da criptografia no direito brasileiro**. São Paulo: Revista dos Tribunais, 2018.

MARTINS, Guilherme M.; LONGHI, João Victor R.; FALEIROS JÚNIOR, José Luiz de M. **Comentários à Lei Geral de Proteção de Dados Pessoais**: Lei 13.709/2018. Induiutaba: Editora Foco, 2022.

MICHAELS, Ralf. The Functional Method of Comparative Law. In: REIMANN, M.; ZIMMERMANN, R. (Orgs.). **The Oxford Handbook of Comparative Law**. 2. ed. Oxford: Oxford Univesity Press, 2019. p. 339–382.

NARAYANAN, A.; SHMATIKOV, V. Myths and fallacies of "personally identifiable information". **Communications of the ACM**, v. 53, n. 6, p. 24, 1 jun. 2010.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. **NISTIR 8053: De-Identification of Personal Information**. 2015.

NEGRI, Sergio M. de C. A.; GIOVANINI, Carolina F. R. Dados não pessoais: a retórica da anonimização no enfrentamento à COVID-19 e o *privacywashing*. **Internet & Sociedade**, v. 1, n. 2, p. 126–149, 2020.

NOUGRÈRES, Ana Brian. Data Protection and Enforcement in Latin America and in Uruguay. *In*: Wright David; DE HERT, Paul (Orgs.). **Enforcing Privacy**: Regulatory, Legal and Technological Approaches. Cham: Springer, 2016. p. 145-180.

PERSONAL DATA PROTECTION COMMISSION. **Guide to basic anonymisation**. Singapore: SG Digital-PDPC, 2022.

PYRGELIS, Apostolos; TRONCOSO, Carmela; DE CRISTOFARO, Emiliano. Measuring Membership Privacy on Aggregate Location Time-Series. **Proceedings of the ACM on Measurement and Analysis of Computing Systems**, v. 2, n. 4, 2020.

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. 40–81, 2018.

ROCHER, L.; HENDRICKX, J. M.; DE MONTJOYE, Y. A. Estimating the success of re-identifications in incomplete datasets using generative models. **Nature Communications**, v. 10, n. 1, 2019.

RUBINSTEIN, Ira S.; HARTZOG, Woodrow. Anonymization and risk. **Washington Law Review**, v. 91, n. 2, p. 703–760, 2016.

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. 1814–189, dez. 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, p. 163-177, 2016.

SWEENEY, Latanya. **Simple Demographics Often Identify People Uniquely**: Data Privacy Working Paper. Pittsburgh: [s.n.], 2000. Disponível em: https://dataprivacylab.org/projects/identifiability/paper1.pdf. Acesso em: 01 mai. 2023.

UNIÃO EUROPEIA. Directiva 95/46/CE do Parlamento Europeu e do Conselho, de 24 de outubro de 1995, relativa à protecção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A31995L0046. Acesso em: 08 mar. 2023.

UNIÃO EUROPEIA. Regulamento (UE) nº 2016/679 do Parlamento Europeu e do Conselho, de 23 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). **Jornal Oficial da União Europeia**, Estrasburgo, 04/05/2016. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679. Acesso em: 08 mar. 2023.

UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia. Grande Secção. **Processo C-582/14, Patrick Breyer v. Bundesrepublik Deutschland**. Luxemburgo, 19 out. 2016. Disponívelem: https://curia.europa.eu/juris/document/document.jsf?docid=184668&doclang=PT. Acesso em: 03 mar. 2023.

UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia. Oitava Secção alargada. **Processo T-557/20, Conselho Único de Resolução v. Autoridade Europeia de Proteção de Dados**. Luxemburgo, 26 abr. 2023. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:62020TJ0557. Acesso em: 09 mai. 2023.

VIMERCATI, Sabrina de C., FORESTI, Sara. Quasi-Identifier. *In*: VAN TILBORG, Henk C. A.; JAJODIA, Sushil (Orgs.). **Encyclopedia of Cryptography and Security**. Springer: Boston, 2011.

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