OBSERVATÓRIO DA LGPD
UNIVERSIDADE DE BRASÍLIA

ANÁLISE COMPARADA
LGPD E GDPR

ORGANIZADORES

LAURA SCHERTEL MENDES

GIOVANNA MILANESE

FELIPE ROCHA DA SILVA

TAYNÁ FROTA DE ARAÚJO

ISABELA MARIA ROSAL

PAULO RICARDO SANTANA

EDUARDA COSTA

ELIS BRAYNER

ANUÁRIO DO OBSERVATÓRIO DA LGPD DA UNIVERSIDADE DE BRASÍLIA

VOLUME 1

Universidade de Brasília
Faculdade de Direito

Anuário do Observatório da LGPD da Universidade de Brasília

Análise comparada entre elementos da LGPD e do GDPR

Volume 1
Brasília-DF
2023

Anuário do Observatório da LGPD da Universidade de Brasília: Análise comparada entre elementos da LGPD e do GDPR.

Para esclarecimentos sobre esta obra, entrar em contato com observatorio.lgpd.unb@gmail.com

Volume 1

Organização
Coordenação Geral: prof.ª Laura Schertel Mendes;
Coordenação Adjunta: Giovanna Milanese;
Coordenação de Pesquisa: Felipe Rocha e Tayná Frota de Araújo;
Revisão e Organização: Eduarda Costa Almeida, Elis Bandeira A. Brayner, Isabela Maria Rosal e Paulo Ricardo da Silva Santana.

Informações

Observatório da LGPD/Unb
Faculdade de Direito
Universidade de Brasília
Campus Universitário Darcy Ribeiro, CEP: 70.910-900, Brasília-DF, Brasil

Dados Internacionais de Catalogação na Publicação (CIP)
(Biblioteca Central da Universidade de Brasília – BCE/UNB)

A636 Anuário do Observatório da LGPD da Universidade de Brasília [recurso eletrônico] : análise comparada entre elementos da LGPD e do GDPR / organização Laura Schertel Mendes … [et al.]. – Brasília : Universidade de Brasília, Faculdade de Direito, 2024. 2 v.

Inclui bibliografia. Modo de acesso: World Wide Web.

ISBN 978-65-00-92398-8 (v. 1).

ISBN 978-65-00-92399-5 (v. 2).

1. Brasil. [Lei geral de proteção de dados pessoais (2018)]. 2. Universidade de Brasília. 3. Proteção de dados. 4. Direito comparado. I. Mendes, Laura Schertel (org.).

CDU 34

AUTORES

Ana Júlia Prezotti Duarte

Andressa Carvalho Pereira

Angélica Opata Vettorazzi

Gabriel de Araújo Oliveira

Gabriel Cabral Furtado

Eduarda Costa Almeida

Fernanda Passos Oppermann Ilzuka

Isabela de Araújo Santos

Júlia Carvalho Soub

Shana Schlottfeldt

Sofia de Medeiros Vergara

Paulo Ricardo da Silva Santana

Rafael Luís Müller Santos

Wanessa Larissa Silva de Araújo

REVISORES

A realização deste anuário contou com a significativa participação de revisores, que atuaram na avaliação e revisão dos artigos submetidos pelos pesquisadores do Observatório, fornecendo orientações e sugestões de melhoria. Oferecemos nosso mais sincero agradecimento pelas valiosas contribuições de cada um.

Ana Luísa Vogado de Oliveira

Angelo Prata de Carvalho

Davi Ory

Gabriel Fonseca

Isabela Maria Rosal Santos

Maria Cristine Lindoso

Matheus Vinicius Aguiar

Paula Baqueiro

Tainá Aguiar Junquilho

Thiago Guimarães Moraes

SUMÁRIO

ANONIMIZAÇÃO DE DADOS PESSOAIS: UM ESTUDO À LUZ DA LGPD E DO RGPD

Ana Júlia Prezotti Duarte[1]

Dispositivos da LGPD	Dispositivos do RGPD
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; **XI** - anonimização: 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; **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. **§ 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. **§ 2º** Poderão ser igualmente considerados como dados pessoais, para os fins desta Lei, aqueles utilizados para formação do perfil comportamental de determinada pessoa natural, se identificada. **§ 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.	**Considerando 26.** Os princípios da proteção de dados deverão aplicar-se a qualquer informação relativa a uma pessoa singular identificada ou identificável. Os dados pessoais que tenham sido pseudonimizados, que possam ser atribuídos a uma pessoa singular mediante a utilização de informações suplementares, deverão ser considerados informações sobre uma pessoa singular identificável. 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 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. Os princípios da proteção de dados não deverão, pois, aplicar-se às informações anónimas, ou seja, às informações que não digam respeito a uma pessoa singular identificada ou identificável nem a dados pessoais tornados de tal modo anónimos que o seu titular não seja ou já não possa ser identificado. O presente regulamento não diz, por isso, respeito ao tratamento dessas informações anónimas, inclusive para fins estatísticos ou de investigação.

[1] Graduanda em Direito na Universidade de Brasília. Integrante do Observatório da LGPD (UnB) e do Grupo de Estudos Constituição, Empresa e Mercado (GECEM/UnB).

Introdução

Este artigo parte de uma investigação bibliográfica acerca do processo de anonimização, em que se busca evidenciar suas implicações quanto à proteção de dados, bem como seus aspectos legais e teóricos. Nesse sentido, por meio de uma análise comparativa entre o artigo 12, da Lei Geral de Proteção de Dados (LGPD), e o Considerando 26, do Regulamento Geral sobre a Proteção de Dados europeu (RGPD), procurar-se-á promover um diálogo entre os dois marcos normativos, destacando os pontos em que convergem ou se distanciam.

Perpassando pelo conceito de dado anonimizado, com o critério da razoabilidade insculpido em lei, tem-se o objetivo de delinear os problemas quanto aos métodos de anonimização, além de investigar tal instituto como processo. Adiante, passa-se a analisar a zona cinzenta entre os conceitos de dados pessoais e dados anonimizados – destacando as soluções encontradas pelo legislador no que se refere ao risco permanente de re-identificação do titular.

Posteriormente, com o fito de gerar maior concretude para o presente estudo e de explorar a visão europeia x visão brasileira acerca do tema em comento, serão analisadas duas decisões jurisprudenciais sobre a problemática, o caso *Breyer*, julgado sob a égide do RGPD, que tratou da definição do dado pessoal, e a Apelação Cível n° 1000631-31.2020.8.26.0452, sob aplicação da LGPD, acerca do tratamento de dados anonimizados no âmbito de um Sistema de Monitoramento Inteligente.

Ao final, verifica-se que a qualificação dos dados depende de uma análise contextual e dinâmica, de acordo com as inovações tecnológicas, não se restringindo à compreensão dicotômica e estática entre dados pessoais e dados anonimizados. Tal assertiva repercute diretamente na atividade de tratamento de dados, demandando do controlador a realização de um teste que leve em consideração todos os meios suscetíveis de serem razoavelmente utilizados. Somente assim, torna-se possível assegurar a conformidade com o regulamento e, ao mesmo tempo, o usufruto das informações, sem que se resulte na violação dos direitos individuais dos titulares.

1. Comentários

1.1. O enigma do dado: a anonimização entre a LGPD e o RGPD

Em ambos os estatutos, o que se extrai dos dispositivos em relevo é que os dados anonimizados, em virtude de não se referirem a uma pessoa natural identificada ou identificável, desde a origem ou após tratamento, não serão considerados dados pessoais – estando fora do seu escopo de aplicação – salvo se for possível descobrir a respectiva autoria (FINKELSTEIN; FINKELSTEIN, 2020). Simplificadamente, se assim caracterizados, os dados podem ser utilizados livremente, não estando sujeitos às restrições impostas pela proteção de dados (BONATTI; KIRRANE, 2019; MACHADO; DONEDA, 2018). Entretanto, essa definição não é trivial; mesmo que identificadores diretos sejam removidos de um banco de dados, ainda será factível reidentificar indivíduos singulares combinando este agregado com outras informações (GRUSCHKA et al., 2018).

A anonimização se trata de um caso "forte" de de-identificação, por meio do qual se busca tornar impraticável, ou até impossível (empregando todos os meios considerados razoáveis) a re-identificação (inclusive pelo próprio controlador) (PINHO, 2017). Salienta-se que a problemática relativa à anonimização envolve os denominados quasi-identificadores, variáveis que podem não identificar sujeitos diretamente, mas que, ocasionalmente, estabelecem uma correlação substancial com elementos identificadores únicos e que podem ser usados para re-identificação indireta (JÚNIOR; MARTINS, 2021).

Os métodos comuns de anonimização existentes são: a supressão, em que os valores de um atributo são completamente removidos ou substituídos por um valor fictício, como um *, sendo aplicável, geralmente, aos identificadores explícitos; a generalização, consistente na modificação da escala ou ordem de magnitude, no caso dos quasi-identificadores, como a data de nascimento (substituindo o formato mês/dia/ano por apenas ano); a permutação, que procura dividir os dados em grupos e embaralhar os valores sensíveis; e, por fim, a perturbação, que diz respeito à substituição de valores removendo o *link* ao dado original, mas de forma a manter suas propriedades estatísticas (BIONI, 2020; GRUSCHKA et al, 2018).

A definição de dado anonimizado prevista no Considerando 26, do RGPD, bem como no artigo 5°, III, e refletida no art. 12, da LGPD, aponta para a existência de três critérios para que o dado seja considerado anônimo: (1) a impossibilidade de se inferir o valor de um atributo de um indivíduo; (2) a inexistência de uma forma conhecida e sistemática de (re)indeitificar os

dados (conhecida como *single-out*); e (3) a incapacidade de conectar dois ou mais registros de uma mesma pessoa (PINHO, 2017; FINCK; PALLAS, 2020; COUVOKIAN; CASTRO, 2014).

Da leitura do artigo 5º, I, da LGPD, evidencia-se que o regulamento brasileiro, assim como o europeu, adotou a estratégia expansionista, consolidando-se como uma legislação de escopo alargado. Nada obstante, diversamente do RGPD, a LGPD não estabeleceu um rol exemplificativo do que pode ser definido como dado pessoal (BIONI; MONTEIRO, 2021).

De um lado, a abordagem reducionista retrai a possibilidade de classificação de um dado como pessoal. Sob esse prisma, somente informações diretamente atreladas a uma pessoa natural *identificada* serão consideradas dados pessoais, a exemplo do RG, do CPF e da biometria. Por outro, conforme se verifica dos contornos da LGPD e do RGPD, com a abordagem expansionista, o vínculo do titular do dado com a informação pode ser mediato, indireto ou inexato e, portanto, se referir a uma pessoa *identificável*, tal como ocorre com profissão, interesses pessoais, endereço de IP e de e-mail corporativo (BIONI; MONTEIRO, 2021).

Outrossim, diferentemente do RGPD, a LGPD dispõe que o dado pode ser considerado pessoal quando empregado para formular perfis comportamentais ("*profiling*") de uma pessoa natural específica e esse tratamento possa culminar na identificação do titular dos dados (KATEIFIDES; MACHADO, 2019; MORIBE et al., 2019). Esse dispositivo da LGDP demonstra um amadurecimento da legislação brasileira no que tange ao regulamento europeu, levando-se em conta o cuidado em regular expressamente as situações oriundas do processamento do *big data* por algoritmos, o que ocorre, comumente, no direcionamento de anúncios publicitários, em matéria de crédito e justiça criminal.

No que concerne ao dado anonimizado, a fim de determinar se esforços razoáveis foram realmente empreendidos, novamente os estatutos se aproximam, como se verá adiante.

Cabe destacar mais uma semelhança entre a legislação brasileira e o regulamento da União Europeia, uma vez que há, no inciso II, do artigo 16 da LGPD, previsão de conservação dos dados, para fins de estudo por órgão de pesquisa, utilizando-se da anonimização, embora esta não seja especificamente recomendada pelo RGPD. Também detém o titular a prerrogativa de requerer a *anonimização*, bloqueio, ou eliminação de dados que sejam desnecessários à finalidade do processamento, ou que estejam sendo submetidos a tratamento em desarmonia com a lei (GRADIM, 2020).

Outra situação possível ocorre quando o *link* entre identificadores explícitos – como o endereço eletrônico ou o número de CPF – e informações sensíveis, a exemplo da orientação sexual e situação financeira de um indivíduo, é o objetivo da análise. Nesta hipótese, a anonimização não é factível e o regulamento deve ser observado. Assim, diante da inviabilidade em tornar o dado anônimo e, simultaneamente, manter a utilidade da informação para fins científicos, estatísticos ou históricos, a pseudoanonimização afigura-se como a técnica mais adequada (GRUSCHKA et al, 2018).

A LGPD brasileira, diversamente do RGPD, não sistematizou adequadamente a figura da pseudoanonimização, muito menos estipulou normativamente incentivos expressos para a sua adoção por parte dos agentes de tratamento de dados. Ao passo que o regulamento europeu estabeleceu até mesmo o afrouxamento de algumas obrigações legais, a LGPD apenas mencionou a pseudoanonimização sem desenvolver propriamente o seu instituto (BIONI, 2020).

De acordo com a definição do artigo 4(9) do RGPD, a pseudonimização consiste no processamento de dados pessoais de modo que não possam mais ser atribuídos a um sujeito específico sem o uso de informação adicional. Nessa perspectiva, é importante consignar que os dados pseudoanonimizados continuam sendo dados pessoais, diferindo-se da anonimização justamente por se referirem a uma pessoa natural identificável. Percebe-se, então, que o Considerando 26 e o seu requerimento de meios razoáveis suscetíveis de serem utilizados permanecem relevantes para a realização desse escrutínio (MOURBY et. al., 2018).

Os autores sugerem que as seguintes perguntas devem ser feitas: a) as pessoas naturais são identificáveis com base na disposição do Considerando 26, tendo em vista todos os meios razoáveis possíveis de serem utilizados? b) se a resposta para a questão acima for afirmativa, a pseudoanonimização foi aplicada com base na disposição do artigo 4(5) do RGPD?

A pseudoanonimização se refere a um processo que reduz o risco de identificação direta, mas que não produz dados anônimos. Dessa forma, dados pseudoanonimizados recaem sob a tutela do regulamento de proteção de dados pessoais. Mas se for preciso reverter este processo, isso pode ser feito pelo controlador, que detém os pseudônimos de mapeamento para os parâmetros identificáveis (GRUSCHKA et al., 2018); como se fosse uma chave capaz de tornar o dado, novamente, pessoalizado.

No entanto, ainda assim é possível chegar à conclusão de que há sim incentivos, mesmo que tácitos, a serem extraídos da LGPD. Na medida em que a pseudoanonimização é o "meio do caminho", a zona de transição entre um dado pessoal e um dado anonimizado, seria possível relacioná-la às diversas referências que a LGPD faz para que os agentes de tratamento "sempre que possível" anonimizem os dados.

Isto pois, a lógica normativa é enxergar o processo de retirada dos identificadores de uma base de dados como uma importante medida de segurança alinhada com o princípio do *privacy by design* proposto na nova legislação. E esse é exatamente o cerne das técnicas de pseudoanonimização, mesmo que a combinação dos dados pseudonimizados com outros conjuntos de dados ainda permita a re-identificação total ou parcial dos indivíduos (BIONI, 2020).

1.2. A faceta escura entre o dado anonimizado e o dado pessoal

No contexto atual, a dificuldade está em avaliar se os métodos disponíveis de anonimização produzem, de fato, dados que são legalmente anônimos. Nessa ótica, é preciso se atentar à questão de que as garantias de tais técnicas e os requerimentos estabelecidos pelas duas leis possuem naturezas distintas. Isso porque, não há como verificar se o conceito legal de anonimização é obedecido em relação a alguns valores padronizados das variáveis *k, l, t e* ε, porquanto o nível de proteção assegurado pela escolha do parâmetro depende das fontes de dados adicionais a que o atacante tenha acesso e essa quantidade de informação é difícil de ser estimada. Por conseguinte, ainda que a discrepância entre as definições técnicas e legais fosse solucionada, o desafio de se mensurar a quantidade de informação disponível seria refletida na escolha imprecisa dos parâmetros aceitáveis (BONATTI; KIRRANE, 2019).

De acordo com Bruno Bioni (2020), o dado anônimo seria a antítese do dado pessoal, ao impedir a identificação da pessoa natural. Elucida-se que a definição de um dado como anônimo tendo por base uma análise contextual que se volta para a suposta irreversibilidade do processo de anonimização traz à tona o problema de seu viés elusivo ou de sua inviabilidade teórica.

Nesta senda, qualquer dado pessoal anonimizado possui o risco inerente de se tornar um dado pessoal, haja vista que sua identificabilidade é remota (identificável) e não imediata (identificada). Por esse motivo, as leis que adotam o conceito expansionista de dados pessoais (que tendem a se *expandir* à medida que a tecnologia o permita) e, concomitantemente, o

estabelecem em franca contraposição aos dados anônimos, teriam grande chance de incidirem em uma redundância normativa.

A fim de manter a coerência, a solução encontrada foi a criação de um filtro que fosse capaz de incorporar a elasticidade desse conceito expansionista, para que fosse possível uma delimitação mais clara da fronteira entre dados pessoais e dados anônimos, sob pena de esta ser sempre transponível (BIONI, 2020).

Dessa maneira, tanto o RGPD quanto a LGPD optaram pelo critério da razoabilidade para definir o espectro do conceito expansionista de dados pessoais. Em outras palavras, o perímetro de elasticidade do conceito de dado pessoal como aquele vinculado a uma pessoa identificável diz respeito ao esforço razoável desprendido no processo de identificação do titular do dado (BIONI, 2020). Esse filtro depende de uma "régua" que enseja a imputação da responsabilidade civil em hipótese de reversão (JUNIOR; MARTINS, 2021).

Consoante Junior e Martins (2021), o conceito de entropia dos dados consegue definir com precisão o espírito da "razoabilidade" insculpido em lei, na medida em que exige elementos mínimos para a confiabilidade da anonimização e para a aferição de seus riscos e falibilidades. Nesse descortino, a entropia atua de modo a indicar o que é preciso para, num corredor repleto de portas fechadas e trancadas com chaves diferentes, cada qual representando as inferências que podem ser feitas, impedir que alguém circule por mais de uma porta, fazendo o cruzamento de diversas informações.

Hodiernamente, verifica-se que não existe mais a pretensão de uma anonimização robusta, tendo se reconhecido amplamente que sempre haverá fatores de risco de identificação e re-identificação de pessoas com o tratamento de dados anonimizados, diante do volume maciço de informações disponibilizadas *online* e do desenvolvimento da capacidade de processamento e análise de algoritmos e de aprendizado de máquina (MACHADO; DONEDA, 2018).

Dessa maneira, torna-se muito mais prudente entender a anonimização e o conceito de dados anonimizados como um processo, mutável e por meio do qual se torna possível manter a utilidade de um banco de dados, e não como um artifício para escapar do regulamento de proteção de dados e de se esvaziar as obrigações que este impõe.

Observa-se que o aspecto mais importante quando se trata da anonimização é a velocidade com que potenciais tecnologias estão se desenvolvendo diuturnamente

(BOLOGNINI; BISTOLFI, 2017), criando um cenário de insegurança tanto para as empresas, que não conseguem assegurar com absoluta certeza a proteção do dado, quanto para os titulares, cujo exercício do direito à privacidade é prejudicado pela cláusula do *take it or leave it* ("pegar ou largar") – pelos altos custos sociais e monetários atrelados a tal liberdade de escolha.

Desse modo, os legisladores brasileiro e europeu tiveram que encontrar uma saída no tocante ao risco de estagnação do conceito legal, considerando-se o desenvolvimento científico e tecnológico. Ao invés de restringi-lo a uma tecnologia que poderia se tornar ultrapassada com o tempo, valeu-se da razoabilidade, conceito este que pode ser constantemente atualizado e ressignificado (BIONI, 2020).

Caberá, então, ao intérprete-aplicador aferir, à luz dos fatos concretos, se é provável que, da conjugação de outros elementos, a identidade do titular dos dados anonimizados possa ser revelada (CORDEIRO, 2018). Para tanto, estabeleceram-se dois eixos de análise.

O primeiro é o objetivo, composto pelos elementos fatoriais: estado da arte da tecnologia, custo e tempo. Cuida-se de uma avaliação dinâmica e circunstancial, que procura evidenciar qual é o grau de investimento financeiro e temporal que deve ser efetuado para se reverter o processo de anonimização. O RGPD também utiliza esses três fatores objetivos para a delimitação da razoabilidade (BIONI, 2020).

O segundo eixo de análise, que está presente apenas na legislação brasileira, é o subjetivo e se centra não nos padrões sociais acerca da reversibilidade de um dado pessoal, mas na capacidade individual de engenharia reversa do agente de tratamento de dados. Além disso, é relevante observar a capacidade subjetiva de terceiros que ingressam no fluxo informacional de uma organização, sobretudo, quando se está diante de atividades em que há enriquecimento de dados que envolvam agentes externos para ensejar uma atividade de tratamento de dados (BIONI, 2020).

No que se refere à criptografia, por exemplo, parcela da doutrina que sustenta ser esta um modo de anonimização parte da premissa de que o dado pessoal criptografado permanece com o mesmo *status* de possibilidade de identificação do titular para o agente que possui a chave criptográfica. Por conseguinte, elegeu-se o eixo de análise subjetivo, centrado na habilidade de engenharia reversa do controlador, e não nos esforços possíveis e razoáveis de qualquer pessoa em obter a informação.

Entretanto, constata-se que, se o sistema não oferecer segurança aos dados cifrados, seja por falha intencional ou eventual, os dados em questão podem ser considerados pessoais. Logo, afigura-se mais consentâneo com a realidade pensar esta espécie de dado como informação pessoal *prima facie*, pseudoanonimizado, sendo aplicável o estatuto de proteção de dados pessoais, mesmo que de forma modulada (MACHADO; DONEDA, 2018).

Enquanto o Considerando 26 do RGPD incorpora um teste baseado no respectivo risco de identificação, o Grupo de Trabalho do Artigo 29o desenvolveu um teste paralelo segundo o qual não pode haver nenhum risco remanescente de identificação para que um dado seja qualificado como anônimo.

Sobre este aspecto, há duas teorias distintas: (i) a teoria relativa, que limita a análise da razoabilidade aos meios e conhecimentos detidos pelo encarregado dos dados; e (ii) a teoria objetiva, inclinada a uma análise abstrata que considera os meios e os conhecimentos detidos não só pelo responsável, mas também por terceiros (CORDEIRO, 2018).

Cordeiro (2018) aponta os principais argumentos que indicam a fragilidade da teoria objetiva: no seu estado mais puro, todos os dados seriam considerados pessoais; a atribuição de importância aos meios e conhecimentos detidos por terceiros impediria que o encarregado dos dados tivesse ciência se está ou não a violar a legislação aplicável, visto que não tem controle sobre a sua anonimidade.

O Considerando 26 do RGPD sugere, inequivocamente, a relevância dos conhecimentos e meios detidos por terceiros. Todavia, não é claro a que terceiros ele alude: todos os sujeitos de direito ou somente uma categoria bem específica, a exemplo de quem trabalha com o responsável dos dados?

Em caso de dúvida, sob a perspectiva da interpretação da lei, a funcionalização exige que se assuma o sentido que mais resguarda os interesses dos titulares de dados pessoais. Da leitura do seguinte trecho fornecido pelo Considerado 26 (“Para determinar se uma pessoa singular é identificável, importa considerar todos os meios suscetíveis de ser razoavelmente utilizados, quer pelo responsável pelo tratamento quer por outra pessoa”), deverão ser considerados os meios detidos por terceiros, mas apenas aqueles *suscetíveis de ser razoavelmente utilizados* (CORDEIRO, 2018).

Por último, faz-se relevante elucidar que podem existir dados os quais, embora totalmente anonimizados, podem, ainda assim, ser considerados dados pessoais. É o que

colaciona a abordagem consequencialista, segundo a qual pouco importa se um tratamento emprega uma informação isolada ou combinada que não se associe direta ou indiretamente a uma pessoa identificada ou identificável. O enfoque está muito mais no impacto que o tratamento pode ter no livre desenvolvimento da personalidade de um grupo ou indivíduo, em que se observa uma zona cinzenta, faceta escura entre os conceitos de dados pessoais e de dados anonimizados (BIONI; MONTEIRO, 2021).

Ganha espaço, então, uma escolha normativa de cunho consequencialista, que levará em conta não só a lógica excludente entre dados pessoais e dados anonimizados, mas também a relação de causa e efeito que uma simples atividade de tratamento de dados pode exercer na vida de seus titulares.

2. Estudos de Caso

2.1. O caso do Sistema de monitoramento de Inteligência: o tratamento de dados anonimizados

O presente caso trata-se do repasse de dados anonimizados ao governo estadual de São Paulo, mediante um acordo de cooperação técnica com empresas de telecomunicações, a fim de mapear os pontos de aglomeração social e, com isso, informar aos cidadãos acerca da incidência da COVID-19. Esta decisão é relevante para o estudo sobre o instituto da anonimização, uma vez que discute o conceito de dado anonimizado, bem como traz uma análise contextual do tratamento dos dados de geolocalização, para, assim, aferir se houve violação do direito à privacidade da Autora. Observe a ementa:

> APELAÇÃO – AÇÃO CONDENATÓRIA – SERVIÇOS DE TELEFONIA MÓVEL – INDEFERIMENTO DA INICIAL – REFORMA – EXISTÊNCIA DE INTERESSE DE AGIR E LEGITIMIDADE PASSIVA – CAUSA MADURA – POSSIBILIDADE DE ADENTRAR NO MÉRITO – REPASSE DE DADOS ANONIMIZADOS AO GOVERNO ESTADUAL – ACORDO DE COOPERAÇÃO TÉCNICA ENTRE EMPRESAS DE TELECOMUNICAÇÃO E GOVERNO ESTADUAL – SISTEMA DE MONITORAMENTO INTELIGENTE (SIMI-SP) – AUSÊNCIA DE VIOLAÇÃO AO DIREITO À PRIVACIDADE – RESPALDO LEGAL, JURISPRUDENCIAL E ADMINISTRATIVO – ENVIO DE MENSAGENS DO GOVERNO À AUTORA INFORMANDO SOBRE AUMENTO DE CASOS DE CORONAVÍRUS (COVID-19) – AUSÊNCIA DE PRÁTICA ABUSIVA – DEFINIÇÃO DE SERVIÇO PARA FINS CONSUMERISTAS – INEXISTÊNCIA DE DANO MORAL

> 1 – Legitimidade passiva da ré para responder por danos advindos de envio de mensagens SMS à autora sem sua autorização. Pertinência subjetiva. 2 – Interesse de agir que não está condicionado à existência de prévio pedido administrativo junto à ré para cancelar o envio de mensagens SMS (inafastabilidade da jurisdição). 3 – Possibilidade de, anulando a r. Sentença por indeferir a inicial incorretamente, julgar o mérito da ação, considerando a teoria da causa madura positivada no atual Código de Processo Civil (CPC, art. 1.013, § 3º, I). Precedentes. 4 – Constitui mero exercício regular do direito o envio de dados anonimizados (informações insuscetíveis de identificação pessoal), tais como dados de geolocalização, às autoridades governamentais, por meio de acordo de cooperação técnica e do Sistema de Monitoramento Inteligente (SIMI-SP). Respaldo legal (Decreto Estadual n. 64.963/20, art. 1º, § único, II; LGPD, art. 5º, III), jurisprudencial (Precedentes recentes do C. STJ e do Órgão Especial deste E. TJSP a respeito disso), e administrativo (Acordo de Cooperação, Pareceres técnicos da AGU e do Ministério da Ciência, e notícias de adoção da mesma medida pela Comunidade Europeia). 5 – Envio pelo Governo do Estado de São Paulo de mensagens SMS informando sobre o aumento de número de casos na região não se qualifica como serviço, à luz do conceito doutrinário e legal (CDC, art. 3º, § 2º), mas pode ser cessado, condenando-se a ré prestadora de serviços à obrigação de não fazer. 6 – Inexistência de violação ao direito à privacidade que evidencia a falta de responsabilidade civil e, portanto, o descabimento de indenização por danos morais. RECURSO PARCIALMENTE PROVIDO (TJSP; Apelação Cível 1000631-31.2020.8.26.0452; Relator (a): Maria Lúcia Pizzotti; Órgão Julgador: 30ª Câmara de Direito Privado; Foro de Piraju - 1ª Vara; Data do Julgamento: 21/09/2020; Data de Registro: 21/09/2020).

Conforme se depreende do acórdão, a parte autora construiu um cenário fantasioso de vigilância social, semelhante àqueles desenvolvidos nas famosas distopias de George Orwell e Aldous Huxley – 1984 e Admirável Mundo Novo, respectivamente. Alegou que a ré estaria repassando informações ao governo estadual, através do chip de seu celular, de modo a violar seu direito fundamental à privacidade, à intimidade e seu direito de ir e vir.

O estado de São Paulo, em meados de abril de 2020, anunciou uma parceria com as operadoras de telefonia com o intuito de executar um Sistema de Monitoramento de Inteligência, destinado à utilização de dados digitais para medir a adesão à quarentena em todo o Estado e também enviar mensagens de alerta para regiões com maior incidência da COVID-19.

Essa parceria deveria observar, em especial, o seguinte comando, previsto no decreto criador do SIMI-SP:

> Art. 1º - Fica instituído o Sistema de Informações e Monitoramento Inteligente – SIMI, consistente em ferramenta de consolidação de dados e informações coligidos por órgãos e entidades da Administração Pública estadual.
> Parágrafo único - O SIMI:
> 1. destina-se a apoiar a formulação e avaliação das ações do Estado de São Paulo para enfrentamento da pandemia da COVID-19;
> 2. **não conterá dados pessoais, assim considerados aqueles relacionados a pessoa natural, identificada ou identificável, limitando-se a dados anonimizados** (grifo nosso).

Nesse sentido, foi destacado na decisão que, desde o início, o governo estadual se preocupou com a captação de dados anonimizados, definidos estes, na LGPD, em seu art. 5º, 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*".* Além disso, fez-se menção à conceituação insculpida no Considerando n° 26, do RGPD, sob a alcunha de *anonymized data*, chamando atenção para a semelhança com a importada pelo ordenamento jurídico brasileiro.

O que se constata de mais relevante na decisão é a análise contextual do tratamento de dados em relevo e a conceituação dos dados anonimizados. Nessa ocasião, o julgador ressaltou a ideia que subjaz o conceito, qual seja, a impessoalidade dos dados enviados, uma vez que irrelevantes para os fins perseguidos pelo acordo, tendo apenas sido concedido ao Estado mecanismos de mapear os pontos de aglomeração social, informação esta imprescindível para o manejo da pandemia do COVID-19. Esclareceu que a remessa é de mera geolocalização impessoal, sem identificação do número de celular, bem como que os dados são agregados, viabilizando a elaboração de gráficos e mapas com os índices de isolamento que serão apresentados na plataforma *Big Data*.

Outro ponto importante diz respeito ao modo de acesso à Plataforma pelo Governo de São Paulo, realizado por meio de login e senha, de maneira que não há compartilhamento de dados pessoais, mas acesso somente a "mapas de calor" e "mapas de identificação de zonas". Vale ressaltar que as informações prestadas pelas operadoras de telefonia não são processadas em tempo real, mas um dia após a conexão, obstando o monitoramento e o reconhecimento da

base de dados da prestadora de telefonia da qual se originaram os dados – a evidenciar que não se instaurou um regime totalitário de violação à privacidade.

A Comunidade Europeia, assim que a crise sanitária eclodiu, veio a público se manifestar precisamente sobre a utilização de dados de geolocalização pelos seus Estados-membros, chegando à mesma conclusão que embasou o governo estadual: a viabilidade de utilização dos dados anonimizados.

No caso brasileiro, em síntese, o Tribunal consignou que nenhum dado pessoal e identificável da autora foi enviado pela ré ao governo estadual, sendo que o repasse de dados anônimos, impessoais e insuscetíveis de pessoalização e identificação não configura violação ao direito à privacidade.

Contudo, vale frisar que o julgador não se desfez da concepção dicotômica e estanque entre dados anonimizados e dados pessoais, pois se valeu da suposição de irreversibilidade do processo de anonimização, o que não encontra respaldo na realidade concreta. Há sempre um risco inerente de re-identificação do titular, diante das ferramentas tecnológicas existentes, bem como da presença de milhares de bancos de dados que guardam informações relevantes sobre os seus titulares. Ainda assim, é possível verificar que houve a observância do parâmetro da razoabilidade previsto no artigo 12, da LGPD, que exige a ponderação dos recursos, do custo e conhecimento necessários para realizar uma re-identificação, com base no contexto tecnológico do momento.

2.2. O caso Breyer: na busca de um parâmetro

No caso Breyer v. Bundesrepublik Deutschland[2], a corte julgou se um endereço IP nas mãos de um controlador deveria ser considerado dado pessoal, quando este não poderia ser usado para identificar um usuário por si só, mas somente quando combinado com dados adicionais coletados do provedor de internet. Neste caso, a corte reputou ser o dado pessoal, pois o controlador tinha meios legais de identificar o usuário, embora esses meios estivessem disponíveis apenas no evento improvável de um ataque cibernético. Ainda que o TJUE não tenha considerado que o dado era anonimizado, a corte sugeriu que as restrições legais ao acesso à chave para re-identificar os dados codificados podem tornar os dados anônimos, ao invés de pseudonimizados, sob certas circunstâncias (PELOQUIN, et al, 2020).

[2]TJUE 19-out.-2016, proc. C-582/14 (Breyer v Bundesrepublik Deutschland).

Conquanto o Parecer 5/2014 do Grupo de Trabalho do Artigo 29o faça referência a "*means likely reasonably to be used"* (meios suscetíveis de serem razoavelmente utilizados, em tradução livre), estabelece-se que todo processo de anonimização deve ser completamente irreversível, alertando contra novas tecnologias que poderiam tornar conjuntos de dados anteriormente presumidos anônimos em dados pessoais (GROOS; VEEN, 2020). Na contramão, no julgamento do caso Breyer, o TJUE adotou uma visão mais flexível, porquanto, mesmo sabendo que um endereço IP dinâmico pode permitir a reidentificação, em algum momento, esta informação não foi suficiente para determinar seu juízo.

Em primeiro lugar, foi observado que não são todos os meios que devem ser levados em conta, mas tão somente aqueles razoáveis e legítimos. Nesse viés, a Corte está mais inclinada a uma abordagem contextual relacionada ao risco, posto que analisa se, concretamente, é possível a reidentificação pelo controlador com o auxílio legítimo de uma terceira parte. De um lado, o Parecer 5/2014 se refere a técnicas abstratas de anonimização que, apenas se forem seguidas à risca, o dado pode ser reputado anonimizado. Por outro, o caso Breyer demanda um teste concreto para o dado em questão e, por conseguinte, para o resultado e o contexto no qual o dado está sendo processado. Desta feita, consoante a decisão, há dois testes distintos (GROOS; VEEN, 2020):

> 1. Quanto ao controlador que não é proibido por lei para realizar a identificação, o dado será anônimo caso este processo exija um esforço desproporcional em termos de tempo, custo e mão-de-obra, de modo que o risco de identificação parece ser, na verdade, insignificante.
> 2. No que se refere ao controlador que não se amolda ao primeiro teste, sendo o risco de identificação significativo, o dado permanecerá anônimo se este processo, seja pelo controlador ou pela ajuda de uma terceira parte, for proibido por lei.

Conforme Groos e Veen (2020), o primeiro teste é contundente sob o ponto de vista legal, mas precisa ser operacionalizado na prática, pois um atacante pode conseguir acesso ao dado, o que constitui uma prática ilegítima em quase todas as jurisdições.

Percebe-se que a decisão do TJUE está em consonância com o teste previsto no Considerando 26, do RGPD, bem como do art. 12, da LGPD, vez que incorpora, essencialmente, uma abordagem baseada no risco para qualificar a informação. Quando há um risco razoável de identificação, o dado deve ser tratado como pessoal, no que concerne a todos os efeitos do regime de proteção de dados. Diversamente, quando o risco for inexpressivo, o

dado pode ser considerado anonimizado, ainda que a identificação não possa ser excluída com absoluta certeza (FINCK; PALLAS, 2020).

De fato, o Parecer 5/2014 traz em seu bojo a ideia de que nenhum risco pode ser tolerado, adotando uma visão muito mais rígida do que aquela extraída do texto legal. Ou seja, ao passo que os dispositivos legais reconhecem que a anonimização nunca pode ser absoluta, haja vista que as tecnologias mudam com o tempo, a postura irrestrita do Grupo de trabalho indica que a anonimização deve ser permanente (FINCK; PALLAS, 2020).

O caso *Breyer* suscita, ainda, outra questão relevante. A ênfase do tribunal na legalidade (apenas em relação ao governo) de obrigar ISPs (provedores de serviço de internet) a revelar os dados necessários para redesenhar um conjunto de dados despersonalizado foi a chave para a sua conclusão. O que nos faz questionar, por um lado, se a ilegalidade de um ato que enseja a identificação justifica que ele seja sempre considerado como razoavelmente improvável. Por outro lado, a adoção de uma abordagem absoluta pode efetivamente descartar a existência de dados anônimos, pois, em última análise, sempre haverá partes capazes de combinar um conjunto de dados com informações adicionais que podem identificá-lo novamente (FINCK; PALLAS, 2020).

Entre as diversas críticas à posição da Corte no caso Breyer, Cordeiro (2018) destaca: a disponibilidade tecnológica, técnica e humana que as empresas possuem; os dados coletados por uma entidade europeia poderem ser repassados a entidades sediadas fora do espaço europeu; as discrepâncias legislativas entre os vários países – o que é lícito sob a ótica do Direito europeu pode ser considerado ilícito em outro ordenamento jurídico; ou a ocorrência recorrente de ataques cibernéticos. O que se defende, portanto, é que, à luz do critério último da razoabilidade, deve-se levar em conta as condutas ilícitas, desde que se possa razoavelmente se dispor delas.

Tal conclusão é reforçada por alguns elementos interpretativos: (i) elemento literal – o termo razoabilidade não exclui ilicitudes; (ii) elemento teleológico – o propósito da lei geral de proteção de dados é resguardar, preventivamente, a devassa dos dados pessoais de sujeitos específicos; e (iii) elemento sistemático – o regime de proteção de dados foi concebido precisamente porque se reconhece que os dados pessoais de cada indivíduo podem ser obtidos ilegalmente (CORDEIRO, 2018).

O autor, ainda, concretiza tal modelo interpretativo à luz de três situações concretas:

1. Qualquer pessoa tem os meios necessários para acessar a informação, em razão de esta ser pública; assim sendo, esses dados não são anônimos, mas pessoais, pois qualquer sujeito possui os meios necessários à sua disposição para desvendar a identidade dos titulares dos dados.

2. O responsável pelo tratamento de dados detém os meios necessários para acessar a informação – situação discutida no acórdão Breyer;

3. Um terceiro tem os meios necessários para acessar a informação.

Aos dois últimos casos é dada solução idêntica – os dados serão anonimizados sempre que o responsável (segunda situação) ou o terceiro (terceira situação) disponham dos meios necessários para identificar os titulares dos dados. Entretanto, nem todos os meios têm relevância jurídica, mas somente os que sejam expectáveis de serem empregados. Diferentemente da posição do TJUE, no acórdão *Breyer*, portanto, o critério legal seria o da razoabilidade e não da razoabilidade mais licitude (CORDEIRO, 2018).

Logo, coadunando com a posição do autor (2018), um meio-termo entre a teoria relativa e a objetiva parece ser a solução mais viável, o que ele denomina de concepção gradual da teoria objetiva, consoante a qual o responsável pelo tratamento dos dados ou as entidades de supervisão devem se ater aos meios detidos por todos os terceiros, porém, limitando a sua análise às informações que, razoavelmente, esses terceiros tenham a seu dispor.

Desta feita, o encarregado do tratamento de dados apenas poderá ser responsabilizado, à luz da boa-fé objetiva, caso esteja dentro do esperado que um terceiro tenha à sua disposição os meios necessários para identificar os titulares dos dados anonimizados e aquele não tenha tomado as precauções devidas pelo regulamento de proteção de dados.

Considerações Finais

Diante da análise comparativa realizada entre o tratamento dado pela LGPD e o RGPD quanto ao instituto da anonimização, percebe-se que, em contraste com a perspectiva legal binária, a realidade opera em um espectro muito mais complexo, tendo em vista o limite móvel que existe entre dados pessoais e dados anônimos, com uma tendência expansiva desses primeiros conforme o avançar da tecnologia.

Nessa linha, nota-se que a qualificação do dado depende do contexto, de forma que a pessoalidade não deve ser vista como propriedade do dado, mas do ambiente no qual este está

inserido. Em síntese, a mesma informação pode ser qualificada como pessoal ou anônima e, consequentemente, sujeitar-se ou não ao regime de proteção de dados.

Ademais, percebe-se que a utilização da razoabilidade como parâmetro extraído dos dispositivos da LGPD e do RGPD (art. 12 e Considerando 26) vai ser avaliada diferentemente se a entidade for uma pessoa com personalidade jurídica privada, uma agência do governo ou uma grande plataforma *online*. Por se tratar de um conceito aberto, torna-se necessário avaliar periodicamente a estratégia de anonimização, considerando o cenário em que o controlador deve assegurar a efetividade de tal processo.

Esse exame é feito de forma *ex ante*, definindo-se o contexto no qual quer se operar e a finalidade do uso do dado anonimizado, com o intuito de compreender quais são os riscos de re-identificação. Ao mesmo tempo, consoante a evolução tecnológica e científica, a técnica pode ser aprimorada *ex post* com meios alternativos de anonimização.

A situação ideal implicaria potencializar concomitantemente a privacidade e a utilidade dos dados, o que, entretanto, é impossível de se alcançar na prática: o desafio advém não apenas das limitações das técnicas existentes, bem como da interconexão entre milhares de bancos de dados contendo informações relevantes. Por isso, afigura-se tão essencial avaliar, concretamente, o fim perseguido com o processamento, para, assim, decidir qual é a técnica mais adequada; afinal, em certas situações, o objetivo do controlador será manter o caráter pessoal dos dados e, em outras, priorizar-se-á a proteção da identidade do sujeito.

Por fim, cumpre adotar uma concepção gradual da teoria objetiva, mais consentânea com a realidade concreta e sem incorrer em absolutismos. Distanciando-se do critério da razoabilidade mais licitude do caso *Breyer* e considerando a dinamicidade existente dentro da qualificação do dado como anonimizado ou pessoal, devem ser levados em conta todos aqueles meios suscetíveis de serem razoavelmente utilizados, para, então, pensar nas possíveis implicações do regime de proteção de dados e nas estratégias que podem ser utilizadas a fim de melhor balancear a privacidade dos titulares e a utilidade das informações.

Referências bibliográficas

BIONI, Bruno. Compreendendo o conceito de anonimização e dado anonimizado. Cadernos Jurídicos, São Paulo, ano 21, nº 53, p. 191-201, 2020.

BIONI, Bruno; MONTEIRO, Ricardo. Data Privacy Brasil. LGPD: O Essencial, 2021. Disponível em: <https://cursos.dataprivacy.com.br/cursos/

exibir/123/combo-ead-muito-alem-da-lgpd> Acesso em: 15 de janeiro de 2021.

BOLOGNINI, Luca; BISTOLFI, Camilla. Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation. Computer Law & Security Review, v. 33, n. 2, p. 171-181, 2017.

BONATTI, Piero A.; KIRRANE, Sabrina. Big Data and Analytics in the age of GDPR. IEEE International Congress on Big Data, 2019.

CAVOUKIAN, Ann; CASTRO, Daniel. Big Data and innovation, setting the record straight: de-identification does work. The Information Technology & Innovation Foundation, Ontario, p. 1-18, jun. 2014.

CORDEIRO, A. Barreto Menezes. Dados pessoais: conceito, extensão e limites. Revista de Direito Civil, Coimbra, v. 3 n. 2, pp. 297-321, 2018.

FINCK, Michèle; PALLAS, Frank. They who must not be identified—distinguishing personal from non-personal data under the GDPR. International Data Privacy Law, 2020.

FINKELSTEIN, Maria Eugenia; FINKELSTEIN, Claudio. Privacidade e lei geral de proteção de dados pessoais. Revista de Direito Brasileira, v. 23, n. 9, p. 284-301, 2020.

GRADIM, Luca Cisneiros. *Análise comparada da lei geral de proteção de dados com o regulamento europeu sobre a proteção de dados e a proteção de dados nos Estados Unidos*. Mestrado em Direito/Relações Internacionais pela FAJS do UniCEUB, 2020.

GROOS, Daniel; VEEN, Evert- Ben van. Anonymized Data and the Rule of Law. European Data Protection Law Review, vol. 6, 2020, p. 498-508.

GRUSCHKA, *et al*. Privacy Issues and Data Protecion on Big Data: A Case Study Analysis under GDPR. IEEE International Congress on Big Data, 2018.

JÚNIOR, José Luiz de Moura Faleiros; MARTINS, Guilherme Magalhães. PROTEÇÃO DE DADOS E ANONIMIZAÇÃO: PERSPECTIVAS À LUZ DA LEI Nº 13.709/2018. REI-REVISTA ESTUDOS INSTITUCIONAIS, v. 7, n. 1, p. 376-397, 2021.

KATEIFIDES, Alexis *et al;* MONTEIRO, Renato *et al*. Comparing privacy laws: GDPR v. LGPD. *One Trust Data Guidance; Baptista Luz Advogados*, 2019.

MACHADO, Diego; DONEDA, Danilo. Proteção de dados pessoais e criptografia: tecnologias criptográficas entre anonimização e pseudonimização de dados. Rev. Trib, v. 998, p. 99-125, 2019.

MOURBY, Miranda *et al*. Are. pseudonymised data always personal data? Implifcations os the GDPR for administrative data research in UK. Computer Law & Security Review, 34, 2018, p. 222-233.

PELOQUIN, David et al. Disruptive and avoidable: GDPR challenges to secondary research uses of data. European Journal of Human Genetics, v. 28, n. 6, p. 697-705, 2020.

PINHO, Frederico. Anonimização de bases de dados empresariais de acordo com a nova Regulamentação Europeia de Proteção de Dados. Universidade do Porto, 2017.


OBSERVATÓRIO DA
LGPD
UNIVERSIDADE DE BRASÍLIA