Considerações iniciais sobre o conceito de dado pessoal no ordenamento jurídico brasileiro

Diego MACHADO*

RESUMO: Trata-se de artigo que se debruça sobre a noção de dado pessoal e seus elementos à luz do ordenamento jurídico brasileiro. O trabalho consiste numa pesquisa teórica de vertente jurídico-dogmática, orientada de acordo com os tipos metodológicos jurídico-compreensivo e jurídico-comparativo. Adotou-se abordagem de acordo com o método dedutivo para se alcançar os objetivos de delimitar os contornos do conceito jurídico de dado pessoal no Brasil e estabelecer parâmetros objetivos a orientar a interpretação-aplicação das normas de proteção de dados pessoais correspondentes. Estruturado em três partes principais, o texto (i) tece considerações sobre as perspectivas regulatórias e abordagens teórico-dogmáticas do conceito de dado pessoal, aproveitando-se de algumas lições relevantes sobre a identificabilidade da pessoa humana; (ii) analisa elementos outros, para além da identificabilidade, relevantes na conceptualização de informação pessoal; e (iii) aprecia a distinção entre dado pessoal e dado não-pessoal na sistemática da LGPD, tendo em conta os confins entre dado anonimizado e dado pseudonimizado, bem como algumas de suas repercussões jurídicas. Conclui-se que a perspectiva expansionista adotada na normativa brasileira se expressa no conceito amplo de dado pessoal, a ensejar não poucas complexidades das abordagens objetiva e relativa no que tange à identificabilidade do titular dos dados, notadamente a partir do critério dos meios suscetíveis de ser razoavelmente utilizados, isto é, dos esforços razoáveis, para a (re)identificação do titular dos dados.

PALAVRAS-CHAVE: Dado pessoal; identificabilidade; dados anonimizados; dados pseudonimizados; LGPD.

TITLE: *Initial Remarks on the Concept of Personal Data under the Brazilian Legal System*

ABSTRACT: *This article focuses on the notion of personal data and its elements in the light of the Brazilian legal system. The work consists of doctrinal qualitative legal research. The methodological approach adopted is in accordance with the deductive method to achieve the aims of outlining the contours of the legal concept of personal data in Brazil, and objective criteria to the interpretation-application of corresponding data protection legal norms. Structured in three main parts, the paper (i) considers regulatory perspectives and theoretical-dogmatic approaches to the concept of personal data, taking advantage of some relevant lessons on the identifiability of the human person; (ii) analyzes elements other than identifiability that are relevant to the conceptualization of personal information; and (iii) appreciates the distinction between personal data and non-personal data in the LGPD system, taking into account the boundaries between anonymized and pseudonymized data, as well as some of its legal effects. It is concluded that the expansionist approach adopted by the Brazilian regulation expresses itself in the broad concept of personal data, creating certain complexities brought by both*

* Professor Adjunto de Direito Civil do Departamento de Direito da Universidade Federal de Viçosa (UFV). Mestre e Doutor em Direito Civil pela Universidade do Estado do Rio de Janeiro (UERJ). Especialista em Privacidade na Autoridade Nacional de Proteção de Dados (ANPD). *Associate Scholar* no projeto *CyberBRICS*. Foi *Fellow Researcher* no *Center for Law, Technology and Society* (CLTS) da Universidade de Ottawa.

DOI: https://doi.org/10.5281/zenodo.18880954 ISSN: 2316-8374

objective and relative approaches as regards to the data subject identifiability, especially the criteria of the means reasonably likely to be used, i.e. of the reasonable efforts, for (re)identifying the data subject.

KEYWORDS: *Personal data; identifiability; anonymised data; pseudonymised data; LGPD.*

1. Introdução

O conceito de dado pessoal é, muito provavelmente, um dos objetos de maiores disputas e discussões na comunidade acadêmica e de profissionais que se debruçam sobre a proteção de dados pessoais.[1] De partida, o próprio termo e sua acolhida sofre variação: não obstante a noção de *dado pessoal* ser adotada na legislação de diversos países, encontra-se também os termos *informação pessoal*[2] e *informação pessoalmente identificável* (“*personally identifiable information*”[3] – *PII*).

A diferença lexical, no entanto, não se traduz necessariamente numa distinção de significado. Da consulta à literatura especializada, percebe-se que os vocábulos “informação” e “dado” não raro se apresentam sobrepostos ou idênticos.[4] Vincenzo Zeno-Zencovich, por exemplo, ao expor o que ele designa de natureza poliédrica da informação, destaca que, numa acepção conteudística, “por informação se entende qualquer dado representativo da realidade que é mantida por um sujeito ou comunicado por um sujeito a outro”.[5] Contudo, essa equivalência tem sido afastada por expressiva parcela dos estudiosos da teoria da informação e suas diversas interfaces.[6] Nesse sentido, já se observou que “o “dado” apresenta conotação um pouco mais primitiva e

[1] Cf. PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, vol. 10, n. 1, p. 40–81, 2018, p. 41.
[2] Na legislação pátria encontra-se essa nomenclatura no art. 4º, IV, da Lei n. 12.527/2011. A mesma opção foi feita pela legislação canadense que regula entidades privadas que tratam informações pessoais – *Personal Information Protection and Electronic Documents Act (PIPEDA)* – e a lei do Estado norte-americano da Califórnia, na recente *California Consumer Privacy Act,* de 2018.
[3] Termo utilizado, principalmente, entre estudiosos da privacidade informacional na América do Norte. Por todos: SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review*, vol. 86, p. 1814–189, dez. 2011; NARAYANAN, Arvind.; SHMATIKOV, Vitaly. Myths and fallacies of personally identifiable information. *Communications of the ACM*, vol. 53, n. 6, p. 24–26, 2010.
[4] MENDES, Laura Schertel. *Privacidade, proteção de dados e defesa do consumidor:* linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014, p. 55.
[5] ZENO-ZENCOVICH, Vincenzo. Informazione (profili civilistici). In: *Digesto* – Sezione Civile, vol. IX. Torino: UTET, 1993, p. 3. Tradução livre de: “per informazione si intende qualsiasi dato rappresentativo della realtà che viene conservato da un soggetto oppure comunicato da un soggetto ad un altro”.
[6] No campo filosófico, a partir da segunda metade do século XX surgem diversas concepções sobre o que é informação: ADRIAANS, Pieter. Information. *The Stanford Encyclopedia of Philosophy* – Edward N. Zalta (ed.). Disponível em: plato.stanford.edu/; PURTOVA, Nadezhda. *Op. cit.*, p. 48-53.

fragmentada".[7] Este, para Raymond Wacks, consiste numa informação em estado potencial, antes de ser transmitida; o dado estaria associado a uma espécie de "pré-informação", anterior à interpretação e ao processo de elaboração. A informação, por sua vez, alude a algo além da representação contida no dado, chegando ao limiar da cognição, e mesmo nos efeitos que esta pode apresentar para o seu receptor. Sem aludir ao significado ou conteúdo em si, na informação já se pressupõe uma fase inicial de depuração de seu conteúdo – daí que a informação carrega em si também um sentido instrumental, apontando à redução de um estado de incerteza.[8]

Deve se considerar, ainda, outra leitura crítica de parte dos estudiosos do tema: no atual cenário de ampla adoção de tecnologias digitais e sistemas algorítmicos, arrisca-se ceder a uma sorte de crença quase religiosa na objetividade dos dados, como se estes fossem apenas a *datificação*[9] da realidade pré-existente em dados brutos (*raw data*).[10] A bem da verdade, afirma-se, a expressão "dados brutos" configura um oximoro, porquanto a prévia seleção de quais são as bases de dados relevantes e os métodos de captura e "datificação" para processamento e criação de conhecimento e modelos algorítmicos a ser ulteriormente aplicados já constitui ato cognitivo de interpretação[11]. Isso resulta em certo adelgaçamento – ou até um abalo – da distinção que se ampara numa visão da informação enquanto *posterius* de ato cognitivo-interpretativo humano, enquanto o dado se constitui num *prius*.

---

[7] DONEDA, Danilo. *Da privacidade à proteção de dados pessoais*. 2. ed. São Paulo: Revista dos Tribunais, 2020, p. 136.

[8] WACKS, Raymond. *Personal Information:* Privacy and the Law. Oxford: Oxford University Press, 1989, p. 25. Vide, também, MENDES, Laura Schertel. *Habeas data* e autodeterminação informativa: os dois lados da mesma moeda. *Direitos Fundamentais & Justiça*, vol. 12, n. 39, p. 185–216, 2018, p. 200; ALBERS, Marion. A complexidade da proteção de dados. *Direitos Fundamentais & Justiça*, vol. 10, n. 35, p. 19–45, 2016, p. 30-32.

[9] Cf. CUKIER, Kenneth; MAYER-SCHÖNBERGER, Viktor. *Big Data:* A Revolution That Will Transform How We Live, Work, and Think. Boston: Houghton Mifflin Harcourt, 2013, p. 73-97.

[10] Cf. HILDEBRANDT, Mireille. *Smart Technologies and The End(s) of Law*. Cheltenham-Northampton: Edward Elgar, 2015, p. xii. Em sentido que parece acolher essa objetividade: Cf. CUKIER, Kenneth; MAYER-SCHÖNBERGER, Viktor. *Big Data:* A Revolution That Will Transform How We Live, Work, and Think. Boston: Houghton Mifflin Harcourt, 2013, p. 73-97.

[11] CRAWFORD, Kate; BOYD, dana. Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, vol. 15, n. 5, p. 662–679, 2012, p. 667-668. Nesta direção: "*However self-contradicting it may be, the phrase raw data — like jumbo shrimp — has understandable appeal. At first glance data are apparently before the fact: they are the starting point for what we know, who we are, and how we communicate. This shared sense of starting with data often leads to an unnoticed assumption that data are transparent, that information is self-evident, the fundamental stuff of truth itself. If we're not careful, in other words, our zeal for more and more data can become a faith in their neutrality and autonomy, their objectivity. Think of the ways people talk and write about data. Data are familiarly 'collected', 'entered', 'compiled', 'stored', 'processed', 'mined', and 'interpreted', Less obvious are the ways in which the final term in this sequence — interpretation — haunts its predecessors. At a certain level the collection and management of data may be said to presuppose interpretation. 'Data [do] not just exist', Lev Manovich explains, they have to be 'generated', Data need to be imagined as data to exist and function as such, and the imagination of data entails an interpretive base*" (GITELMAN, Lisa; JACKSON, Virginia. Introduction. In: GITELMAN, Lisa (Org.). *'Raw data' is an oxymoron*. Cambridge; London: The MIT Press, 2013, p. 2-3).

Pois bem. Dito isso, para fins da análise e delimitação do conceito de dado pessoal que ora se propõe, os termos *dado* e *informação* são considerados sinônimos, como o é, inclusive, na legislação do Brasil.[12] Desta forma, ambas nomenclaturas – *dado pessoal* e *informação pessoal* – serão utilizadas de forma intercambiável.

A compreensão do conceito juridicamente positivado de dado pessoal é tarefa imprescindível para a interpretação do alcance normativo de diversas leis de proteção de dados atualmente em vigor mundo afora. A título de exemplo, o *Children's Online Privacy Protection Ac*t (COPPA) de 1998, estatuto norte-americano de proteção da infância no uso da internet, é aplicável a todos que coletam informação pessoal de menores de 13 anos, estabelecendo critérios para o legítimo tratamento desses dados.[13] Pode-se citar também o GDPR e a Convenção 108, do Conselho da Europa. Nesses dois textos legislativos do modelo europeu de proteção de dados, o conceito de dado pessoal[14] também é chave para discernir o âmbito material de aplicação do regime inscrito nos instrumentos normativos.[15]

Para explicar a importância e função dogmática da noção – e instigar importante reflexão –, Ian Kerr usou como metáfora a figura de um porteiro, extraída da conhecida (e enigmática) parábola "Diante da Lei" contada na prosa de Franz Kafka, no seu romance "O Processo".[16] Assim como o porteiro se coloca como guardião da porta de entrada da

---

[12] Basta a leitura dos artigos 5º, I, da Lei Geral de Proteção de Dados – LGPD (Lei n. 13.709, de 14 de agosto de 2018), e 4º, IV, da Lei de Acesso à Informação – LAI (Lei n. 12.527, de 18 de novembro de 2011).

[13] Conforme previsto na *Section 1303, a, 1,* o COPPA estatui: "(1) IN GENERAL. — It is unlawful for an operator of a website or online service directed to children, or any operator that has actual knowledge that it is collecting personal information from a child, to collect personal information from a child in a manner that violates the regulations prescribed under subsection (b)" (ESTADOS UNIDOS. *Children's Online Privacy Protection Act of 1998*. Disponível em: law.cornell.edu/).

[14] Conforme art. 4, 1, do Regulamento n. 2016/679 da UE (GDPR), dado pessoal "*means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person*". Já a Convenção 108, cujo texto foi modernizado em 2018, define no art. 2, *a*, dado pessoal como "*any information relating to an identified or identifiable individual ('data subject')*".

[15] HOOFNAGLE, Chris J.; VAN DER SLOOT, Bart; BORGESIUS, Frederik Z. The European Union general data protection regulation: what it is and what it means. *Information & Communications Technology Law*, vol. 28, n. 1, p. 65–98, 2019, p. 72.

[16] No trecho inicial da parábola narrada ao personagem Josef K. se lê: "Diante da lei há um porteiro. Um homem do campo dirige-se a este porteiro e pede para entrar na lei. Mas o porteiro diz que agora não pode permitir-lhe a entrada. O homem do campo reflete e depois pergunta se então não pode entrar mais tarde. "É possível", diz o porteiro, "mas agora não", Uma vez que a porta da lei continua como sempre aberta, e o porteiro se põe de lado, o homem se inclina para olhar o interior através da porta. Quando nota isso, o porteiro ri e diz: "Se o atrai tanto, tente entrar apesar da minha proibição. Mas veja bem: eu sou poderoso. Eu sou apenas o último dos porteiros. De sala para sala, porém, existem porteiros cada um mais poderoso do que o outro. Nem mesmo eu posso suportar a visão do terceiro", O homem do campo não esperava tais dificuldades: a lei deve ser acessível a todos e a qualquer hora, pensa ele; agora, no entanto, ao examinar mais de perto o porteiro, com seu casaco de pele, o grande nariz pontudo e a longa barba tártara, rala e preta, ele decide que é melhor aguardar até receber a permissão de entrada. O porteiro lhe dá um banquinho e deixa-o sentar-se ao lado da porta. Ali fica sentado dias e anos. Ele faz muitas tentativas para ser admitido, e cansa o porteiro com os seus pedidos" (KAFKA, Franz. *O Processo*. Trad. e posfácio de Modesto Carone. 1. ed. São Paulo: Companhia das Letras, 2005, p. 214).

lei e aparenta ser impenetrável – não se sabe ao certo –, o conceito de dado pessoal e sua interpretação parece ser o canal, de diâmetro algo indeterminado e maleável, pelo qual se acessa o regime de proteção de dados pessoais com os direitos e garantias nele contidos.[17]

Na doutrina, duras críticas já foram endereçadas a essa centralidade do conceito de dado pessoal, a exemplo do que faz Paul Ohm, que, considerando problemática a noção "sempre em expansão" de "informação pessoalmente identificável", sustenta o abandono do conceito como pilar da proteção jurídica da privacidade informacional[18-19]. Não é essa a perspectiva adotada neste artigo científico. Muito embora o debate crítico seja essencial, o abandono da formal distinção de dado pessoal e dado não pessoal pelas leis de proteção de dados, possivelmente terá efeito oposto: todo dado será considerado de caráter pessoal[20], alargando sobremaneira a aplicabilidade dos regimes de proteção de dados, de sorte a resultar num processo de agravada erosão de sua efetividade.

Desenvolve-se no presente trabalho uma pesquisa teórica de vertente jurídico-dogmática, orientada de acordo com os tipos metodológicos jurídico-compreensivo e jurídico-comparativo,[21] porquanto na atividade de interpretação do conceito jurídico lança-se mão do procedimento analítico de decomposição dos elementos constitutivos do conceito, integrado da confrontação da noção com outras categorias jurídicas relevantes, a saber, dados anonimizados e dados pseudonimizados. Adota-se, ainda, uma abordagem segundo o método dedutivo a fim de se alcançar os objetivos de delimitar os confins da noção de dado pessoal no ordenamento jurídico brasileiro e estabelecer parâmetros objetivos a orientar a interpretação-aplicação das normas de proteção de

---

[17] KERR, Ian. Foreword. In: GRATTON, Éloïse. *Understanding Personal Information: Managing Privacy Risks*. Markham: LexisNexis, 2013, p. iii.

[18] Afirma o autor: "*At the very least, we must abandon the pervasively held idea that we can protect privacy by simply removing personally identifiable information (PII). This is now a discredited approach. Even if we continue to follow it in marginal, special cases, we must chart a new course in general*" (OHM, Paul. Broken promises of privacy: responding to the surprising failure of anonymization. *UCLA Law Review*, n. 57, p. 1701-1777, 2010, p. 1742).

[19] Há que se salientar que não há aqui qualquer tomada de posição que considera sinônimos os conceitos de privacidade e de proteção de dados pessoais. Haja vista, no entanto, a profundidade de análise que o tema requer e os limites deste artigo científico, anota-se apenas que, ao longo deste trabalho, optou-se pelo uso do termo "privacidade informacional" apenas quando associado a um contexto da tradição jurídica de países da América do Norte (em sua doutrina, legislação e jurisprudência). Isso se explica pelo fato de que a partir da década de 1970, em termos de política legislativa e regulatória, o discurso e os debates sobre o regramento da atividade de tratamento de dados pessoais se consolidaram em torno dos conceitos de "privacidade informacional" (*information privacy*), principalmente nos EUA, e "proteção de dados" (*data protection*), entre países europeus especialmente. BENNETT, Colin J.; RAAB, Charles D. *The Governance of Privacy*: Policy Instruments in Global Perspective. 1. ed. London-New York: Routledge, 2003, p. 16.

[20] Cf. PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, vol. 10, n. 1, p. 40–81, 2018, p. 80; WACHTER, Sandra; MITTELSTADT, Brent D. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, vol. 2019, n. 2, p. 494-620, 2019, p. 578.

[21] 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.

dados de pessoais atinentes ao conceito jurídico. Para se desincumbir da investigação nos termos desenhados, imprescindíveis foram a coleta e análise de conteúdo (i) bibliográfico, de textos nacionais e estrangeiros, com teor jurídico, filosófico e de áreas afins à ciência da computação; (ii) legislativo e (iii) jurisprudencial, a abranger tanto o sistema jurídico do Brasil como outras experiências, especialmente a da União Europeia (UE).

O artigo está estruturado em três partes principais. No item 2, inicia-se as considerações sobre as perspectivas regulatórias e abordagens teórico-dogmáticas do conceito de dado pessoal, aproveitando-se de algumas lições relevantes sobre a identificabilidade da pessoa humana, especialmente da experiência jurídica da União Europeia, para o âmbito do sistema jurídico pátrio e da LGPD. Já no item 3, elementos outros, para além da identificabilidade, são analisados para a conceptualização de informação pessoal, terminando-se, no item 4, por apreciar a distinção entre dado pessoal e dado não pessoal na sistemática da LGPD, tendo em conta os confins entre dado anonimizado e dado pseudonimizado, bem como algumas de suas repercussões jurídicas.

2. Dado pessoal: perspectivas e abordagens

Numa primeira aproximação do objeto da conceituação, pode-se depreender a distinção de abordagens de técnica legislativa utilizadas para a construção dos conceitos restrito e amplo, ou que observa, correspondentemente, o que Daniel Solove e Paul Schwartz nomeiam de *perspectivas reducionista* e *expansionista* de política regulatória de proteção de dados.[22] Na *conceitualização restrita*, por dado pessoal entende-se a representação de fatos sobre pessoa identificada, isto é, representação referente a alguém que se conhece e individualiza em meio a certo grupo ou coletividade.

O processo de identificação é possível a partir de elementos informativos chamados identificadores, "os quais mantém relação particularmente privilegiada e próxima com certo indivíduo".[23] Os identificadores podem ser, por sua vez, *diretos* ou *indiretos*. O típico identificador direto de um indivíduo é o seu nome completo. Constituído por prenome e sobrenome, o nome da pessoa humana é o primeiro sinal distintivo da

[22] SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review*, vol. 86, p. 1814–189, dez. 2011, p. 1871-1877. A respeito do tema, vide também, BIONI, Bruno. *Xeque-Mate:* o tripé de proteção de dados pessoais no xadrez das iniciativas legislativas no Brasil. São Paulo: GPOPAI, 2015.

[23] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 12. Disponível em: ec.europa.eu/. Tradução livre de: "[...] *which hold a particularly privileged and close relationship with the particular individual*".

individualidade, forma de identificação social da pessoa entre os demais membros de dada comunidade.[24] Endereço eletrônico e número do CPF também podem ser considerados identificadores diretos. No entanto, um identificador direto pode não ser bastante pois pode se configurar, a título de exemplo, a homonímia, de maneira que identificadores indiretos (ou "quase-identificadores"[25]) como a nacionalidade, a raça, a filiação ou o CEP da residência, e mesmo características fenotípicas, podem ser necessários para se distinguir alguém. A categoria dos identificadores indiretos é, assim, conexa ao "fenômeno das 'combinações únicas'".[26]

Uma vez adotada, essa concepção pode ser implementada *por específica tipificação*[27] em lei de quais identificadores (diretos ou indiretos) são reputados informação pessoal. Isto é, se o dado se enquadrar dentro de uma das categorias-tipo elencadas pelo legislador, ele se torna dado pessoal por obra da lei. Fora da moldura legal não haveria, então, dado pessoal, nem se aplicaria o pertinente regime de proteção de dados pessoais. No ordenamento jurídico dos EUA, o conceito restrito é tipificado em dois importantes diplomas do *statutory law*: o *Privacy Act*, de 1974, e o já mencionado COPPA.[28]

De outro lado, o *conceito amplo* estende seu alcance para além da pessoa natural identificada: também é informação de caráter pessoal aquela relativa a pessoa *identificável*. Há dado pessoal não apenas quando houver a presença de identificadores

[24] O nome atende nesse caso ao que Lara Trucco denomina etapa de atribuição do processo de "identificação pessoal jurídica" (TRUCCO, Lara. *Introduzione allo studio dell'identità individuale nell'ordinamento costituzionale italiano*. Torino: Giappichelli, 2004, p. 4-5). Vide, também, ARTICLE 29 DATA PROTECTION WORKING PARTY. *Op. cit.*, p. 13. Ressalta-se, todavia, que o nome da pessoa humana é aqui tratado tão somente sob o prisma da identificação, não contemplando a profundidade própria da perspectiva da tutela da personalidade.

[25] Em linguagem utilizada por áreas como a ciência da computação e segurança da informação, fala-se em *"quasi-identifiers"*: "A *quasi-identifier* is a set of attributes that, in combination, can be linked with external information to reidentify (or reduce uncertainty about) all or some of the respondents to whom information refers" (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).

[26] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Op. cit.*, p. 13. No julgamento do RE 673.707/MG, pelo STF, o Min. Luiz Fux sinalizou 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).

[27] "*In the context of the specific-types approach, if information falls into an enumerated category, it becomes per se PII by operation of the statute*" (SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review*, vol. 86, p. 1814–189, dez. 2011, p. 1831-1832).

[28] O *U. S. Code, Title 15, Chapter 9, § 6501 (8)*, determina: "*Personal information. The term 'personal information' means individually identifiable information about an individual collected online, including — (A) a first and last name; (B) a home or other physical address including street name and name of a city or town; (C) an e-mail address; (D) a telephone number; (E) a Social Security number; (F) any other identifier that the Commission determines permits the physical or online contacting of a specific individual; or (G) information concerning the child or the parents of that child that the website collects online from the child and combines with an identifier described in this paragraph*",

diretos ou indiretos que diferem precisamente um indivíduo. Os dados que *potencialmente* conduzem à individuação da pessoa são igualmente tomados como informação pessoal. É esta noção que se vê inscrita na normativa da EU,[29] que influenciou, por sua vez, diversos sistemas como os da Argentina[30] e da Colômbia.[31]

Existem dados ou identificadores que, apesar de não individuarem efetivamente alguém, caso tratados com técnicas que são razoavelmente acessíveis e/ou em conjunto com dados suplementares, podem levar à identificação de seu titular. Ainda que o agente responsável pelo tratamento dos dados não possa identificar com precisão a pessoa natural a quem se referem as informações processadas, com algum esforço ele pode se valer de meios disponíveis para a obtenção dos dados adicionais aptos a fazê-lo. Considere-se como exemplo os dados de geolocalização obtidos de usuários de *smartphones*. O agente de tratamento que, num arco temporal, coleta apenas dados de geolocalização a partir de *software* instalado em dispositivo móvel, ainda que não trate dados como nome ou outro identificador direto ou indireto, processa informação pessoal. Os pontos de localização registrados no curso do tempo formam um único padrão de deslocamento individual que possibilitam a identificação do usuário sem que isso requeira desmesurados esforços ou complexo recurso a grande quantidade de dados auxiliares.[32]

Em dezembro de 2019, jornalistas do *The New York Times* receberam de fonte anônima uma base de dados com aproximadamente 50 bilhões de pontos de localização registrados de telefones celulares, entre vários meses de 2016 e 2017, de mais de 12 milhões de pessoas que rotineiramente trafegavam em cidades como Washington, Nova Iorque, São Francisco e Los Angeles. Com poucas fontes suplementares, os jornalistas conseguiram identificar várias pessoas, entre militares, oficiais de segurança pública e políticos. Participantes de protesto político foram também rastreados em ocasião que se deu nos primeiros dias da posse de Donald Trump na presidência dos EUA em 2017: de uma multidão de aproximadamente 500.000 mil manifestantes que desciam as ruas da

[29] Vide a redação do art. 4, 1, GDPR, e art. 2, *a*, Convenção 108+, reproduzidos na nota de rodapé nº 15.
[30] *Ley Estatutaria 1581/2012*, art. 3º, *c: "Dato personal: Cualquier información vinculada o que pueda asociarse a una o varias personas naturales determinadas o determinables*".
[31] *Ley 25.326/2000*, art. 2º, 1: "*Datos personales: Información de cualquier tipo referida a personas físicas o de existencia ideal determinadas o determinables*".
[32] Ao analisar 15 meses de dados anonimizados de telefonia móvel de cerca de 1,5 milhão de usuários, pesquisadores do MIT e da *Universite Catholique de Louvain*, na Bélgica, descobriram que foram necessários poucos dados para identificar exclusivamente 95% dos usuários. Partindo das atualizações horárias da localização de um usuário, rastreadas por *pings* de seu *smartphone* nas torres de celular mais próximas, enquanto se moviam, ou realizavam chamadas e trocavam mensagens de texto, os pesquisadores puderam identificar o indivíduo com apenas quatro tipos de dados (*"data points")*. Usando somente dois tipos de dados, eles conseguiram identificar cerca de 50% dos usuários (DE MONTJOYE, Yves-Alexandre et al. Unique in the Crowd: The privacy bounds of human mobility. *Scientific Reports*, vol. 3, p. 1–5, 2013).

capital, foi possível seguir o rastro de pessoas meses antes e depois do protesto, precisando inclusive os locais de trabalho e residência. O conjunto de dados analisados foi originariamente coletado por empresa de dados de geolocalização sem que houvesse a associação desses dados a identificadores como nome e endereço eletrônico.[33] O caso mostra, desse modo, que quem trata dados de geolocalização, ainda que sem processar outros identificadores, processa informação de pessoas identificáveis com razoável esforço próprio.[34]

Na concepção ampla de dados pessoais, diferentes abordagens podem ser escolhidas a depender da importância que se atribui à atuação e esforços de terceiros na avaliação da identificabilidade da pessoa humana[35]. Denomina-se *relativa* a orientação que tem em vista apenas os esforços e meios próprios do controlador que trata os dados em questão, ou seja, é abordagem que, na apreciação da identificabilidade, desconsidera a atividade e esforços de outrem.[36] De outra parte, a perspectiva que considera o potencial de identificação não só por meios e esforços do responsável pela operação de tratamento mas também de qualquer outra pessoa ou ente, é chamada de *absoluta* ou *objetiva*.

No caso acima mencionado sobre processamento de dados de geolocalização, ambas abordagens chegam à mesma conclusão: há informação pessoal; com a atividade de terceiros (*v. g.*, jornalistas) ou não, a identidade dos usuários de *smartphones* pode ser

[33] THOMPSON, Stuart A.; WARZEL, Charlie. The Privacy Project - Twelve Million Phones, One Dataset, Zero Privacy. *The New York Times*, 19 dez. 2019. Disponível em: nytimes.com/.

[34] Sendo um dos entrevistados para a matéria do *The New York Times*, Paul Ohm afirma que "[r]eally precise, longitudinal geolocation information is absolutely impossible to anonymize" (THOMPSON, Stuart A.; WARZEL, Charlie. *Op. cit.*). Para uma melhor compreensão os dados de geolocalização, as formas de coleta e quem os acessa, vide: GRAY, Stacey. A Closer Look at Location Data: Privacy and Pandemics. Future of Privacy Forum, 2020. Disponível em: fpf.org/.

[35] 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*, vol. 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*, vol. 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*, vol. 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*, vol. 10, n. 1, p. 11–36, 2020, p. 17.

[36] 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: search.coe.int/.

obtida.[37] Modifique-se, contudo, o exemplo para o tratamento de registros de endereço IP.[38] O provedor de aplicação de internet que não coleta dados cadastrais[39] dos seus usuários, mas processa dados de tráfego e armazena registros de endereço IP de terminais que acessaram seu sítio eletrônico, não consegue identificá-los de forma imediata. Para fazê-lo, terá de algum modo acessar os registros de conexão e os dados cadastrais armazenados pelos respectivos provedores de conexão à internet. Partindo do ângulo *relativo*, o endereço IP não configura dado pessoal para o provedor de aplicação de internet, já que, por não tratar dados cadastrais, não está apto a identificar os usuários. Mas adotando-se a abordagem *absoluta* ou *objetiva*, o endereço IP é visto como dado pessoal pois há a possibilidade de determinação da identidade junto a provedores de conexão[40] – tendo em consideração os esforços de terceiros, portanto.

Esta última abordagem parece ter inspirado a interpretação empreendida na UE, tanto pelo Grupo de Trabalho de Proteção de Dados do Artigo 29[41] como pelo Tribunal de Justiça da União Europeia (TJUE) nos casos C-70/10, *Scarlet Extended SA v. Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)*, e C-582/14, *Patrick*

[37] Sobre dados de geolocalização de dispositivos móveis, o Superior Tribunal de Justiça, no RMS 61.302/RJ, proferiu acórdão reconhecendo sua qualificação como dados pessoais. Na oportunidade, o tribunal apreciou recurso interposto pelo Google contra decisão judicial que ordenou a entrega de "dados estáticos (registros), relacionados à identificação de usuários em determinada localização geográfica que, de alguma forma, possam ter algum ponto em comum com os fatos objeto de investigação por crimes de homicídio" (BRASIL. Superior Tribunal de Justiça (Terceira Seção). Recurso Ordinário em Mandado de Segurança n. 61.302/RJ. Relator: Min. Rogério Schietti Cruz. Brasília, 26/08/2020. *Diário de Justiça eletrônico*, Brasília, DF, 04/09/2020). Sobre o caso em questão, vide CASO Marielle: Google deve quebrar sigilo e fornecer dados para investigação. *Migalhas*, 26 ago. 2020.

[38] O *Internet Protocol* consiste em protocolo essencial para o funcionamento da internet, pois permite o roteamento de pacotes de dados e a comunicação entre dispositivos. Para tanto, necessário é especificar origem e destino desses pacotes de dados mediante mecanismo de endereçamento: o endereço IP. A título ilustrativo, o conjunto de números 400.456.12.789 seria um endereço IP a individuar uma máquina conectada à internet.

[39] De acordo com o art. 10, § 3º, do Marco Civil da Internet (MCI), e o art. 11, § 2º, do Decreto n. 8.771/2016, são considerados *dados cadastrais* (i) a filiação, (ii) o endereço, e (iii) a qualificação pessoal, entendida como nome, prenome, estado civil e profissão do usuário.

[40] Considere, no ordenamento jurídico brasileiro, o que dispõe o art. 22 do MCI: "Art. 22. A parte interessada poderá, com o propósito de formar conjunto probatório em processo judicial cível ou penal, em caráter incidental ou autônomo, requerer ao juiz que ordene ao responsável pela guarda o fornecimento de registros de conexão ou de registros de acesso a aplicações de internet. Parágrafo único. Sem prejuízo dos demais requisitos legais, o requerimento deverá conter, sob pena de inadmissibilidade: I - fundados indícios da ocorrência do ilícito; II - justificativa motivada da utilidade dos registros solicitados para fins de investigação ou instrução probatória; e III - período ao qual se referem os registros",

[41] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Privacy on the Internet – An integrated EU Approach to On-line Data Protection*. Bruxelas: [s. n.], 2000. Disponível em: ec.europa.eu/.

Breyer v. Bundesrepublik Deutschland.[42] No primeiro julgado, a menção à caracterização do endereço IP como dado pessoal é feito em sede de *obiter dictum*, mas no caso *Breyer* compõe de fato as razões de decidir (*ratio decidendi*) do acórdão a qualificação do endereço IP (dinâmico) como informação relativa a pessoa natural identificada ou identificável.[43]

A análise do potencial de identificação de certa informação pelo agente de tratamento ou por outro sujeito – conforme a orientação objetiva – não pode ser feita em abstrato apenas, como se bastasse uma possibilidade puramente hipotética.[44] No direito europeu, desde a Diretiva 95/46/CE vigora o *critério dos meios suscetíveis de ser razoavelmente utilizados*, que busca ser balizado por fatores objetivos como custos e tempo de trabalho exigidos para a identificação, o estado da arte da tecnologia existente no período de duração do tratamento e os riscos de falhas técnicas, por exemplo.[45]

Trata-se de critério dependente de aspectos contextuais – *v. g.*, estágio de desenvolvimento das tecnologias de rastreamento de comportamento *online* e de algoritmos de aprendizado de máquina – que, em última análise, faz da caracterização

---

[42] Neste caso, o TJUE afirmou nos fundamentos da decisão: "Na medida em que esse considerando [n. 26 da Diretiva 95/46] faz referência aos meios suscetíveis de serem razoavelmente utilizados quer pelo responsável pelo tratamento quer por «qualquer outra pessoa», a sua redação sugere que, para que um dado possa ser qualificado de «dado pessoal» na aceção do artigo 2.º, alínea a), da referida diretiva, não é necessário que todas as informações que permitem identificar a pessoa em causa tenham de estar na posse de uma única pessoa",
"O facto de as informações suplementares necessárias para identificar o utilizador de um sítio Internet não serem detidas pelo prestador de serviços de meios de comunicação em linha, mas pelo fornecedor de acesso à Internet desse utilizador, não parece, assim, suscetível de excluir que os endereços IP dinâmicos registados pelo prestador de serviços de meios de comunicação em linha constituam, para ele, dados pessoais na aceção do artigo 2.º, alínea a), da Diretiva 95/46". Ressalte-se, porém, que o TJUE não adotou na prática uma abordagem puramente objetiva ou absoluta, porquanto inovou ao exigir que os meios razoavelmente suscetíveis de ser utilizados pelo agente ou provedor de serviço devem ser meios *lícitos* necessariamente. Cf. 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*, vol. 7, p. 163-177, 2016, p. 167-168; CORDEIRO, António Barreto Menezes. Dados pessoais: conceito, extensão e limites. *Revista de Direito Civil*, Coimbra, vol. 3, n. 2, p. 297-321, 2018, p. 321; PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, vol. 10, n. 1, p. 40–81, 2018, p. 64.

[43] Sobre a discussão do endereço IP como dado pessoal, vide: 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*, vol. 32, n. 2, p. 256–271, 2016, p. 7-10; BORGESIUS, Frederik Zuiderveen. The Breyer Case of the Court of Justice of the European Union: IP Addresses and the Personal Data Definition. *European Data Protection Law Review*, vol. 3, n. 1, p. 130–137, 2017; MACHADO, Diego C.; SOUZA, Carlos Affonso P. Tutela da privacidade, guarda de registros e portas lógicas no direito brasileiro. In: FERRARI, Isabela; BECKER, Daniel (Org.). *Regulação 4.0* – Novas tecnologias sob a perspectiva regulatória. São Paulo: Revista dos Tribunais, 2019, p. 247-277.

[44] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 15.

[45] O Considerando 26 do GDPR estabelece que: "*The principles of data protection should apply to any information concerning an identified or identifiable natural person.* [...] *To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments*".

como dado pessoal um estado *dinâmico*.[46] Esse caráter maleável do conceito, se de um lado contribui para que uma lei de proteção de dados acompanhe e possa se adaptar às transformações sociotécnicas, por outro pode dar ensejo a insegurança em alguma medida, uma vez que os avanços das tecnologias orientadas por dados (*data-driven technologies*) se desenvolvem em passo acelerado e poderiam conferir caráter pessoal a dado antes considerado anonimizado, devido à disponibilização de novas tecnologias capazes de realizar esta reidentificação.[47] Maiores são os benefícios do que as desvantagens, todavia.

Junto à importância dada ao contexto do tratamento da informação como parâmetro para a interpretação do conceito de dado pessoal, surgem proposições para a leitura desta noção com alicerce em análise de gestão de risco.[48] Na conceptualização formulada por Daniel Solove e Paul Schwartz, os autores sugerem um modelo em que se avalia o *risco de identificabilidade*, tendo em consideração os meios suscetíveis de ser razoavelmente utilizados para identificação.[49] Afirmam que o modelo aventado apresenta a informação num *continuum* "que começa com a ausência de risco de identificação em uma extremidade, e termina com pessoas identificadas na outra"[50] – é dizer, segue-se a seguinte divisão: (i) informação não identificável, (ii) informação identificável e (iii) informação identificada. Éloïse Gratton elabora, por sua vez, um esquema hermenêutico

[46] PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, vol. 10, n. 1, p. 40–81, 2018, p. 47. Cf. ARTICLE 29 DATA PROTECTION WORKING PARTY. ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 15; EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS. *Handbook on European Data Protection Law*. Luxembourg: Publications Office of the European Union, 2018, p. 88; FINCK, Michèle; PALLAS, Frank. They Who Must Not Be Identified - Distinguishing Personal from Non-Personal Data Under the GDPR. *International Data Privacy Law*, vol. 10, n. 1, p. 11–36, 2020, p. 12.

[47] Sobre esse apontamento crítico, afirma Purtova: "*The resulting standard of the reasonable likelihood of identification is quite broad and context-dependent, leading to one major consequence: the status of data as 'personal' is dynamic, i.e. the same dataset may not obviously be personally identifiable at the start of processing, or from the perspective of the controller, given the tools and data available to him, but become, or appear to have been all along, identifiable from the perspective of another person or once the circumstances change*" (PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, vol. 10, n. 1, p. 40–81, 2018, p. 47).

[48] Henry Pearce afirma que, em termos gerais, abordagens regulatórias baseadas na gestão do risco "*can be described as regulatory strategies that involve the targeting of enforcement and resources on the basis of assessments of the risks that a particular regulated activity poses to the regulator's objectives. The key components of these assessments will be the evaluations of the risks of noncompliance and calculations pertaining to the impact that said noncompliance may have on the regulatory body's ability to achieve its objectives. In its idealised form, therefore, risk management-based regulation offers an evidence-based means of targeting the use of resources and of prioritising attention to the highest risks in accordance with a transparent, systematic, and defensible framework*" (PEARCE, Henry. Big data and the reform of the European data protection framework: an overview of potential concerns associated with proposals for risk management-based approaches to the concept of personal data. *Information and Communications Technology Law*, vol. 26, n. 3, p. 312–335, 2017, p. 321).

[49] SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review*, vol. 86, p. 1814–189, dez. 2011, p. 1878.

[50] SCHWARTZ, Paul M.; SOLOVE, Daniel J. *Op. cit.*, p. 1877.

com lastro no *risco de dano* que certas operações de tratamento de dados geram,[51] desconsiderando como dado pessoal a informação cujo tratamento implica baixo risco.

Apesar da importância dessas abordagens num cenário em que as leis de proteção de dados pessoais cada vez mais lançam mão de instrumentos regulatórios orientados pelo controle dos riscos da atividade de tratamento de informações pessoais (*v. g.*, relatório de impacto à proteção de dados, regime de responsabilidade civil),[52] são grandes os desafios a serem enfrentados para a implementação prática de efetiva gestão de riscos numa sociedade orientada por dados. Dentre os problemas pontuados por Henry Pearce, destaca-se aqueles relacionados à concepção do modelo de avaliação dos possíveis danos decorrentes da atividade de tratamento e a probabilidade de sua ocorrência. São dois os pontos críticos: (i) complexidade de se *calcular* o nível de risco inerente à sistemas algorítmicos de análise de dados (e que se baseiam em aprendizado de máquina); e (ii) a probabilidade de sua real ocorrência.[53] As concepções acima citadas de Daniel Solove, Paul Schwartz, e Éloïse Gratton, por exemplo, não descem ao detalhe do efetivo cálculo dos riscos sob análise – se é que tal é sequer possível.

É seguro afirmar que o direito brasileiro seguiu a orientação consagrada na legislação e jurisprudência aplicadas na UE, adotando a noção ampla de dado pessoal. Como pode se verificar no artigo 5º, I, da LGPD, dado pessoal no sistema brasileiro de proteção de dados é "informação relacionada a pessoa natural identificada ou identificável". Este conceito legal, na verdade, endossa o que a Lei de Acesso à Informação prescreve, desde 2011, no seu artigo 4º, IV, pelo qual a informação pessoal é "aquela relacionada à pessoa natural identificada ou identificável".

3. Elementos conceituais para além da identificabilidade

Importante observação a se fazer sobre a análise até aqui feita sobre a definição de dado pessoal é a de que apenas um dos seus elementos constitutivos foi objeto de apreciação: a *identificabilidade*. Para um exame pormenorizado, outros elementos devem, porém, ser estudados, o que se passa a fazer segundo a *perspectiva analítica*, que, guardando

[51] GRATTON, Éloïse. *Understanding Personal Information: Managing Privacy Risks*. Markham: LexisNexis, 2013, p. 223 et seq.
[52] Sobre a disciplina do tratamento de dados pessoais centrada numa abordagem sobre o risco, vide MANTELERO, Alessandro. Responsabilità e rischio nel Reg. UE 2016/679. *Le Nuove Leggi Civili Commentate*, vol. XL, n. 1, p. 144-164, 2017, p. 147 et seq.
[53] PEARCE, Henry. Big data and the reform of the European data protection framework: an overview of potential concerns associated with proposals for risk management-based approaches to the concept of personal data. *Information and Communications Technology Law*, vol. 26, n. 3, p. 312–335, 2017, p. 324-325.

certa sintonia ao ordenamento jurídico brasileiro, foi empregada pelo extinto Grupo de Trabalho do Artigo 29 em parecer sobre o conceito de dado pessoal datado de 2007. De acordo com o documento opinativo, em referência à então vigente Diretiva 95/46/CE, quatro são os elementos constitutivos da definição: (i) qualquer informação; (ii) relativa a; (iii) pessoa natural; (iv) identificada ou identificável.[54] Trata-se de abordagem que, mesmo não possuindo efeito vinculante na esfera da UE, se tornou influente, direta ou indiretamente, para a atuação de instituições europeias que integram o sistema de proteção de dados pessoais do bloco regional[55]. A jurisprudência do TJUE dá sinais de ter encampado essa mesma perspectiva, ainda que não o reconhecendo expressamente.

Já havendo explorado acima o item *iv*, passa-se, pois, à análise dos elementos *i, ii* e *iii*.

Diferentemente do GDPR[56] e da lei de proteção de dados da Colômbia,[57] que aludem a *"qualquer informação"*, ou da lei argentina,[58] que se refere a "*informação de qualquer tipo*", a redação legal da LGPD faz menção do termo *"informação"* somente. Tal como se faz noutras jurisdições, a interpretação da lei brasileira deve ser a mais abrangente possível, a fim de compreender qualquer informação passível de identificar alguém.

"[O] conceito de dado pessoal inclui qualquer espécie de afirmação sobre uma pessoa",[59] portanto. Neste sentido, a medida em que a informação é acessível ou divulgada ao público,[60] é dizer, "pública" ou "privada", não a desnatura como informação pessoal.

---

[54] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 6.

[55] PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, vol. 10, n. 1, p. 40–81, 2018, p. 59 et seq.; KUNER, Christopher; BYGRAVE, Lee A.; DOCKSEY, Christopher. *The EU General Data Protection Regulation:* A Commentary. 1. ed. Oxford: Oxford University Press, 2020, p. 109-111.

[56] GDPR, art. 4º, 1.

[57] *Ley Estatutaria 1581/2012*, art. 3º, *c:* "Dato personal: Cualquier información vinculada o que pueda asociarse a una o varias personas naturales determinadas o determinables".

[58] *Ley 25.326/2000*, art. 2º, 1: "Datos personales: Información de cualquier tipo referida a personas físicas o de existencia ideal determinadas o determinables".

[59] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 6. Tradução livre de: "[...] the concept of personal data includes any sort of statements about a person".

[60] A LGPD dispõe no art. 7º, §§ 3º e 4º, sobre o regime protetivo aplicável a informações acessíveis ao público, tais como as informações processuais publicadas em diários oficiais: "Art. 7º [...] § 3º O tratamento de dados pessoais cujo acesso é público deve considerar a finalidade, a boa-fé e o interesse público que justificaram sua disponibilização. § 4º É dispensada a exigência do consentimento previsto no caput deste artigo para os dados tornados manifestamente públicos pelo titular, resguardados os direitos do titular e os princípios previstos nesta Lei", Guardadas as devidas proporções, a assertiva se alinha aos fundamentos do acórdão do STJ no Recurso Especial n. 1.758.799/MG, da relatoria da Min. Nancy Andrighi: "[...] o fato de alguém publicar em rede social uma informação de caráter pessoal não implica o consentimento, aos usuários que acessam o conteúdo, de utilização de seus dados para qualquer outra finalidade, ainda mais com fins lucrativos" (BRASIL. Superior Tribunal de Justiça (Terceira Turma). Recurso Especial n. 1.758.799/MG. Relatora: Min. Nancy Andrighi. Brasília, 12/11/2019. *Diário de Justiça eletrônico*, Brasília, DF, 19/11/2019). Sobre a proteção de dados pessoais e sua compatibilização com o princípio da publicidade dos atos processuais no âmbito do Poder Judiciário, vide: CUEVA, Ricardo Villas Bôas. Proteção de dados pessoais no Judiciário. *Revista do Advogado*, São Paulo, n. 144, p. 6–12, nov./2019.

Aliás, tal se evidencia com clareza na LGPD a partir da expressa disciplina aplicada a operações de tratamento envolvendo dados de acesso público ou publicizados pelo próprio titular.[61] Também não afeta a qualificação da informação pessoal o ser verdadeira ou incorreta,[62] até porque em havendo incorreção, o regime de proteção de dados pessoais prevê remédios para a retificação dos dados.[63] De igual modo, é dado pessoal tanto a informação de ordem subjetiva como a objetiva. Em outras palavras, a informação sobre alguém que se traduz em opinião ou avaliação a seu respeito, bem como a que não possui esse fator subjetivo, configura dado pessoal.

No caso C-434/16, *Peter Nowak v. Data Protection Commissioner,* o TJUE, resolvendo questão processual que lhe foi submetida, se manifestou sobre a qualificação como dado pessoal de respostas escritas formuladas por um candidato em exame profissional e das respectivas anotações lançadas pelo examinador. Na decisão, a corte europeia fundamentou que:

> [...] o emprego da expressão «qualquer informação» no âmbito da definição do conceito de «dado pessoal», constante do artigo 2º, alínea a), da Diretiva 95/46, reflete o objetivo do legislador da União de atribuir um sentido amplo a esse conceito, que não está limitado às informações sensíveis ou de ordem privada, mas engloba potencialmente qualquer tipo de informações, tanto objetivas como subjetivas sob forma de opiniões ou de apreciações [...].[64]

O caso remete, ainda, à abrangência do conceito jurídico frente a diferença de origem dos dados, pois que os dados consubstanciados na avaliação feita por examinador são dados derivados da apreciação crítica das respostas formuladas pelo candidato para aferição do seu nível de conhecimento ou aprendizagem. À luz do critério da origem, os dados podem ser classificados em fornecidos, observados, derivados e inferidos.[65]

Dados fornecidos (*individually provided data*) são dados que se originam diretamente de ação, em geral, voluntária e consciente da pessoa humana, o titular dos dados. É o que

[61] LGPD, art. 7º, §§ 3º, 4º, 6º e 7º.
[62] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 6.
[63] LGPD, art. 18, II.
[64] UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia. Segunda Seção. *Case C-434/16, Peter Nowak v. Data Protection Commissioner*. Luxemburgo, 20 dez. 2017. O tribunal interpretou a Diretiva 95/46 no caso, mas também fez expressa alusão às disposições do GDPR.
[65] Cf. WACHTER, Sandra; MITTELSTADT, Brendt D. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review,* vol. 2019, n. 2, p. 494–620, 2019, p. 516 et seq.; ARTICLE 29 WORKING PARTY. *Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679*. Brussels: [s.n.], 2018, p. 8.; INFORMATION COMMISSIONER'S OFFICE. *Big data, artificial intelligence, machine learning and data protection*. [s.l: s.n.], p. 12-13. Disponível em: ico.org.uk/.

ocorre, por exemplo, com as fotos publicadas em conta de rede social, a comunicação do próprio endereço eletrônico, dados cadastrais registrados em formulários de sítios eletrônicos. Dados pessoais de crianças cujo tratamento é fundado no consentimento dos pais ou responsáveis pelo menor, também se enquadram nessa categoria. Já os *dados observados* (*observed data*) consistem em informações objeto da observação por terceiros e capturados em formato digital. Esses dados podem ser registrados no momento de sua criação ou transmitidos para um intermediário após a observação.[66] A coleta de dados observados se dá muito provavelmente sem a consciência do titular dos dados sobre a operação de tratamento; há, de certa forma, o fornecimento indireto ou passivo dos dados pelo titular.[67] Se ajustam a essa categoria, por exemplo, os dados do histórico de navegação em *websites*, a caligrafia, dados sobre uso de jogos de videogame, dados obtidos pelo emprego de câmeras CCTV com programas de reconhecimento facial, dados de geolocalização, e dados sobre a temperatura corporal. O aumento exponencial da coleta de dados observados possui relação com a multiplicidade de sensores usados para reconhecimento de atividade, e aplicações de internet e os diversos provedores que atuam como intermediários.[68]

Por fim, os *dados derivados ou inferidos* (*derived or inferred data*) são dados resultantes de outros dados pessoais ou não-pessoais, sejam eles fornecidos ou observados, devido a raciocínio ou operações lógico-matemáticas não probabilísticas[69] (*v. g.*, deriva-se o país de residência do indivíduo a partir do seu CEP) ou em razão da aplicação de modelos estatísticos complexos baseados em algoritmos de mineração de dados e sistemas de aprendizado de máquina[70] (*v. g.*, escore de crédito). A inferência

---

[66] ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking* – Summary of the OECD Privacy Expert Roundtable, 2014, p. 5.

[67] WACHTER, Sandra; MITTELSTADT, Brendt D. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, vol. 2019, n. 2, p. 494–620, 2019, p. 516.

[68] WORLD ECONOMIC FORUM. *Rethinking personal data*: Trust and Context in User-Centred Data Ecosystems. [s.l: s.n.], p. 16.

[69] ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking* – Summary of the OECD Privacy Expert Roundtable, 2014, p. 5.

[70] De acordo com relatório do Fórum Econômico Mundial, essa categoria de dados surge com os novos sistemas algorítmicos: "*Advanced computational analytics and machine learning create a third category of data that is 'inferred' and synthesized from an array of different data types (including data directly related to individuals and data that is not connected to them). Inferred data is generally more of an amalgam of different originating data types and is generally used for predictive purposes*" (WORLD ECONOMIC FORUM. *Op. cit.*, p. 16-17). De forma mais detalhada, porém, afirma relatário da OCDE: "*From a historical perspective, it was pointed out that most of the data categories noted above have been in existence for a long time. 'Inferred data', however, was said to have more recent origins, situated in the early 1980s (when companies first began to develop credit risk scores). There has, however, been a remarkable increase in both the volume and variety of available data sets. In particular, there has been substantial growth in the amount of 'observed', 'derived' and 'inferred' data. This development has been attributed to a number of factors. One is the continuous decrease in data storage and communication costs. A second is the Internet, which has led to a significant expansion in the types and amount of data being created since the mid-1990s. Finally, a third important development has been the proliferation, particularly in the 21st century, of new sensor technologies, which allow for detailed observations in both physical and digital environments*" (ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Op. cit.*, p. 5-6).

computacional de dados é absolutamente fundamental às tecnologias orientadas por dados e sistemas algorítmicos. Modelos ou perfis são inferidos, e então formados, a partir do reconhecimento de padrões em bases de dados comportamentais fornecidos e/ou observados (*v. g.,* modelo computacional de *credit scoring* para avaliação de risco de inadimplemento).

Foi mediante a inferência de dados que se deu a criação e aplicação dos modelos do *Office of Qualifications and Examinations Regulation* (*Ofqual*) para avaliação de estudantes no Reino Unido em 2020. Devido às restrições de locomoção para o combate à pandemia de COVID-19 e consequente suspensão das aulas e impossibilidade de realização de provas presenciais, os exames *GCSE, AS, A-level, Extended Project Qualification and Advanced Extension Award,* foram feitos com um método diferente. Primeiramente, os professores escolares foram incumbidos de (i) atribuir a cada estudante, por cada disciplina, a provável nota que obteria caso as aulas tivessem continuado e a prova realizada, e (ii) estabelecer um *ranking* comparativo com os demais alunos da mesma escola que obtiveram igual nota.[71] Esses dados foram, então, tratados por sistema algorítmico cujo modelo matemático também processava os resultados agregados da respectiva escola, em cada matéria, nos três anos anteriores. Isto é, a autoridade britânica partiu do entendimento de que as notas do ano de 2020 seriam consistentes com a performance das escolas nos anos precedentes.[72] Na efetiva aplicação do referido modelo ou perfil aos candidatos, dos dados pessoais tratados na entrada (*input*) e processados pelo sistema, infere-se outros dados, visto que da identificação e representação do titular como adequado a um perfil ou categoria resulta na conclusão de novas informações,[73] tais como a medida das competências e habilidades cognitivas do estudante.

Apesar de originado da atividade do agente de tratamento, esse aspecto do dado inferido não parece capaz *per se* de desnaturar um dado como informação pessoal. Na LGPD, aliás, o disposto nos arts. 11, § 1º e 13, § 1º, convergem com essa assertiva. De um lado, prescreve-se que o agravado regime protetivo dos dados sensíveis também se aplica a qualquer operação de tratamento de dados pessoais não sensíveis que *revele dados pessoais sensíveis*.[74] Por exemplo, considere-se, a partir do contexto de período de

[71] OFQUAL. *Executive Summary – Awarding GCSE, AS, A level, advanced extension awards and extended project qualifications in summer 2020:* interim report. Londres: [s. n.], 13 ago. 2020. Disponível em: assets.publishing.service.gov.uk/.
[72] A-LEVELS and GCSEs: How did the exam algorithm work? *BBC,* [s.l.], 20 ago. 2020. Disponível em: bbc.com/.
[73] De semelhante forma poder-se-ia pensar sobre o escore de crédito atribuído a certo consumidor.
[74] De acordo com o art. 5º, II, da LGPD, dado pessoal sensível é "dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural".

entrada no Brasil e do país de origem, a derivação de informação sobre pertencimento a minoria étnica imigrante,[75] ou a inferência computacional do estado de gravidez de consumidora a partir de seu histórico de compras.[76] Já o art. 13, § 1º, proíbe a divulgação dos resultados ou de qualquer excerto de estudo ou de pesquisa em saúde pública que *revele dados pessoais*. Numa e noutra disposição normativa, enfim, o objeto versa sobre dados pessoais derivados ou inferidos.[77]

O segundo elemento conceitual (*"relativo a"*) estabelece a necessidade de a informação possuir vínculo com uma pessoa. Esse ponto de ligação existe, a princípio, quando a informação é sobre uma pessoa identificada ou identificável.[78] Isso, no entanto, não significa que dados que remetem imediatamente a objeto ou evento não possam constituir informação pessoal. É possível que, de maneira indireta, esses dados também se relacionem com indivíduos.[79] Por exemplo, a partir do uso da Interface de Programação de Aplicativos (API) *HTML5 Battery Status,* o nível de bateria de um *smartphone* pode ser processado como dado identificador para rastrear usuários *web* em curtos intervalos de tempo.[80] Ressalte-se, todavia, que neste específico exemplo, o critério dos meios suscetíveis de ser razoavelmente utilizados muito provavelmente importaria num obstáculo ao reconhecimento do caráter pessoal da informação.

O liame que faz da informação *relativa a* pessoa natural se forma em razão do *conteúdo*, *finalidade* ou *resultado* do tratamento do dado, seja este um nexo direto ou indireto. Não se trata de critérios que devem ser cumulativamente observados, porém; são aspectos passíveis de articulação *alternativa* a se considerar na interpretação das circunstâncias do caso.[81]

---

[75] Cf. INTERNATIONAL COMMITTEE OF THE RED CROSS. *Handbook on Data Protection in Humanitarian Action*. 2. ed. Geneva: ICRC, 2020, p. 71.
[76] DUHIGG, Charles. How Companies Learn Your Secrets. *The New York Times Magazine,* 16 fev. 2012. Disponível em: nytimes.com/.
[77] Para estudo de maior fôlego a respeito dos dados pessoais sensíveis inferidos, cf. MACHADO, Diego; MENDES, Laura Schertel. A proteção dos dados sensíveis inferidos: um comentário ao Caso C-184/20 do Tribunal de Justiça Europeu. *Revista de Direito do Consumidor*, São Paulo, vol. 144, a. 31, p. 97-121, nov.-dez./2022.
[78] *"In general terms, information can be considered to "relate" to an individual when it is about that individual*" (ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 9).
[79] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 53-54.
[80] Confira o estudo: OLEJNIK, Lukasz; ACAR, Gunes; CASTELLUCCIA, Claude; DIAZ, Claudia. *The leaking battery A privacy analysis of the HTML5 Battery Status API*. Disponível em: eprint.iacr.org/.
[81] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 10-12; PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, vol. 10, n. 1, p. 40–81, 2018, p. 54. Dado o caráter não cumulativo e não procedimental, não parece correta a descrição do esquema de análise proposto pelo Grupo de Trabalho do Artigo 29 como um "*three-step model*", tal como sugere Wachter e Mittelstadt (WACHTER, Sandra; MITTELSTADT, Brent D. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, vol. 2019, n. 2, p. 494-620, 2019, p. 517-519).

A informação se relaciona com a pessoa a partir do conteúdo quando diz respeito à vida de alguém – *e. g.*, número de registro geral de identidade, características físicas, débitos fiscais, fotografia, histórico escolar. Já o critério da finalidade se realiza quando a informação é usada, ou há probabilidade de ser usada, "com o *propósito* de avaliar, tratar de certa forma ou influenciar o estado ou o comportamento de uma pessoa".[82] O uso de *cookies*[83] em sede de publicidade comportamental, por exemplo, é feito com o objetivo de monitorar hábitos de navegação dos usuários, ou proporcionar-lhes uma experiência personalizada de navegação em sítios eletrônicos e outras aplicações de internet.[84] Por fim, há informação relativa a alguém se "é provável que seu uso tenha *impacto* nos direitos e interesses de certa pessoa, levando em consideração todas as circunstâncias em torno do caso".[85] No já citado processo *Peter Nowak v. Data Protection Commissioner*, o TJUE analisou as respostas escritas por candidato de exame profissional a partir do critério do resultado, inclusive, ao que concluiu que o uso dessa informação "é suscetível de ter um efeito sobre os direitos e interesses dele ou dela, na medida em que pode determinar ou influenciar, por exemplo, a chance de ingressar na profissão almejada ou de obter o cargo pretendido".[86]

A interpretação sob as lentes do critério do *resultado* do tratamento sobre direitos e interesses juridicamente protegidos da pessoa identificada ou identificável se mostra especialmente relevante em operações que envolvem sistemas algorítmicos e aplicações de aprendizado de máquina,[87] inclusive quando houver tomada de decisão automatizada. Há aqui, entretanto, importantes alertas a serem feitos em relação a dois perigos: (i) de

---

[82] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Op. cit.*, p. 10. Tradução livre de: "*with the purpose to evaluate, treat in a certain way or influence the status or behaviour of an individual*".

[83] *Cookies* são arquivos de texto que, uma vez enviados por *websites* e armazenados nos computadores dos usuários a partir do navegador utilizado, funcionam como identificadores eletrônicos.

[84] Este exemplo, em da verdade, também atende ao critério do conteúdo. Cf. ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 2/2010 on online behavioral advertising*. Bruxelas: [s. n.], 2010, p. 9. Disponível em: ec.europa.eu/; 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*, vol. 32, n. 2, p. 256–271, 2016, p. 5-6.

[85] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 11. Tradução livre de: *"their use is likely to have an impact on a certain person's rights and interests, taking into account all the circumstances surrounding the precise case"*. António Barreto Menezes Cordeiro parece fazer uma leitura diversa desse critério. De acordo com o autor português, "como resultado entende-se toda a informação que não incida sobre uma pessoa (conteúdo) e que não vise avaliá-la ou influenciá-la (finalidade), mas que, em abstrato, o permita fazer" (CORDEIRO, António Barreto Menezes. Dados pessoais: conceito, extensão e limites. *Revista de Direito Civil*, Coimbra, vol. 3, n. 2, p. 297-321, 2018, p. 305).

[86] Tradução livre de: *"is liable to have an effect on his or her rights and interests, in that it may determine or influence, for example, the chance of entering the profession aspired to or of obtaining the post sought*" (UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia. Segunda Seção. *Case C-434/16, Peter Nowak v. Data Protection Commissioner*. Luxemburgo, 20 dez. 2017).

[87] Cf. WACHTER, Sandra; MITTELSTADT, Brent D. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, vol. 2019, n. 2, p. 494-620, 2019, p. 518.

todos dados ao cabo se tornarem dados pessoais; e (ii) de se embrenhar nas inconsistências de análise consequencialista do conceito de dado pessoal.

Como bem indica Nadezhda Purtova, no cenário de atual difusão das tecnologias digitais orientadas por dados, o uso do referido critério que leva em conta o impacto do tratamento de dados sobre a esfera de direitos e interesses juridicamente tutelados de pessoa natural pode conduzir à excessiva e incontida ampliação da noção de dado pessoal.[88] Para fundamentar a assertiva, a autora lança mão do exemplo dos dados sobre as condições do tempo de certa localidade (em Eindhoven), que poderá, sim, ser reputado de caráter pessoal em caso de iniciativa de cidade inteligente (projeto *Stratumseind 2.0*) que, combinando os dados sobre as condições de tempo com outros gerados por sensores distribuídos ao longo de determinada rua (vídeo de transeuntes, dados sonoros e sobre qualidade do ar, quantidade de pedestres por dia/hora, uso de rede Wi-Fi etc.), podem acarretar impactos em direitos e interesses de indivíduos. Fosse relevante os dados sobre as condições de tempo e chuva para gerir rede inteligente de transporte público, poder-se-ia falar, então, em dados pessoais, eis que poderiam afetar o exercício da liberdade de locomoção de alguém.

Outro caminho que enseja problemas expressivos é o de se valer de chave de leitura consequencialista[89] para com ela ler o critério do resultado e, por conseguinte, compreender se se trata de informação "relativa a" pessoa identificada ou identificável. Nessa linha, defendendo que a estrutura regulatória da LGPD abre espaço "para uma escolha normativa *consequencialista*",[90] há parcela da doutrina brasileira que afirma que a noção de dado pessoal abarca todas "as hipóteses nas quais o tratamento de dados pode ocasionar efeitos negativos sobre uma pessoa ou um grupo de pessoas".[91] A visão consequencialista da conceituação de dado pessoal toma aportes do pragmatismo jurídico, de sorte que sua aplicação pode esbarrar em limitações e insuficiências que aparecem na construção de uma teoria pragmatista da privacidade informacional (ou proteção de dados), como a de Daniel Solove. O método tripartite de raciocínio deste jurista estadunidense[92] implica que ninguém tem direito à privacidade *ex ante*. As pessoas têm direito à proteção da privacidade apenas se um dano ou lesão à privacidade

[88] PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, vol. 10, n. 1, p. 40–81, 2018, p. 57 et seq.
[89] Sobre o consequencialismo, cf. BRANDÃO, Rodrigo; FARAH, André. Consequencialismo no Supremo Tribunal Federal: uma solução pela não surpresa. *Revista de Investigações Constitucionais*, vol. 7, n. 3, p. 831–858, Curitiba, set.-dez./2020, p. 835 et seq.
[90] BIONI, Bruno R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019, p. 78. Grifos no original.
[91] BIONI, Bruno R. *Op. cit.*, p. 82.
[92] Cf. SOLOVE, Daniel J. *Understanding Privacy*. Cambridge; London: Harvard University Press, 2008, p. 91 et seq.

for verificado, isto é, aos indivíduos não se reconhece a titularidade do direito à privacidade na eventualidade de um interesse colidente sobrepujar seu interesse de tutela à privacidade em termos de benefícios sociais ou coletivos. Decerto, essa não é uma narrativa adequada para sistemas jurídicos cujas normas constitucionais reconhecem à toda pessoa humana o direito fundamental à privacidade e à proteção de dados pessoais.[93]

Ademais, ao partir de um contexto específico (contextualismo[94]) advogando "a primazia das consequências na interpretação",[95] a referida abordagem consequencialista pode gerar uma hermenêutica reducionista, nas palavras de Luciano Floridi: bem pensadas as coisas, se se considera dado pessoal o que pode ocasionar efeitos negativos sobre uma pessoa ou um grupo de pessoas, não havendo consequências reputadas socialmente negativas ou indesejáveis do tratamento de dados, não há se falar em interesses juridicamente relevantes a proteger,[96] a exemplo do acesso à informação sobre a lógica de funcionamento de sistema de aprendizado de máquina que gera baixo risco a direitos fundamentais.

No tocante ao elemento do item *iii*, tem-se que, para ter caráter pessoal, as informações devem ser relativas a *pessoa natural*. Este componente do conceito possui laço estreito com a teleologia das leis de proteção de dados pessoais direcionada à tutela e realização de direitos e liberdades fundamentais da pessoa humana.[97] O primeiro desdobramento

---

[93] Em semelhante linha argumentativa, Ronald Dworkin critica o pragmatismo como uma teoria cética sobre os direitos: "Segundo nossa apresentação abstrata, "conceitual", da prática jurídica, uma pessoa tem a pretensão juridicamente protegida de ganhar um processo se esse direito decorrer de decisões políticas anteriores. O convencionalismo oferece uma teoria positiva, não cética, dos direitos que as pessoas possuem: elas têm como pretensões juridicamente asseguradas todos os direitos que as convenções jurídicas extraem de decisões políticas tomadas no passado. O direito como completeza é também uma teoria não cética das pretensões juridicamente protegidas: sustenta que as pessoas têm como pretensões juridicamente protegidas todos os direitos que são patrocinados pelos princípios que proporcionam a melhor justificativa da prática jurídica como um todo. O pragmatismo, ao contrário, nega que as pessoas tenham quaisquer direitos; adota o ponto de vista de que elas nunca terão direito àquilo que seria pior para a comunidade apenas porque alguma legislação assim o estabeleceu, ou porque uma longa fileira de juízes decidiu que outras pessoas tinham tal direito" (DWORKIN, Ronald. *O império do direito*. Trad. Jefferson Luiz Camargo. São Paulo: Martins Fontes, 1999, p. 186).

[94] BUTLER, Brian E. Legal Pragmatism: Banal or Beneficial as a Jurisprudential Position? *Essays in Philosophy*, vol. 3, n. 2, p. 269-286, jun./2002, p. 278.

[95] POSNER, Richard A. *Overcoming Law*. Cambridge: Harvard University Press, 1995, p. 252. Tradução livre de: "the primacy of consequences in interpretation".

[96] Luciano Floridi critica o que ele denomina de "interpretação reducionista" da privacidade informacional por se caracterizar pelo consequencialismo, isto é, que define a privacidade informacional a partir de indesejáveis efeitos: *"The advantage of the ontological interpretation over the reductionist one is then that consequentialist concerns may override respect for informational privacy, whereas the ontological interpretation, by equating its protection to the protection of personal identity, considers it a fundamental and inalienable right, so that, by default, the presumption should always be in favour of its respect. As we shall see, this is not to say that informational privacy is never negotiable in any degree*" (FLORIDI, Luciano. The ontological interpretation of informational privacy. *Ethics and Information Technology*, vol. 7, n. 4, p. 185–200, 2005, p. 195).

[97] De acordo com o art. 1º da LGPD, é objetivo da lei "proteger os direitos fundamentais de liberdade e de privacidade [...] da pessoa natural".

lógico dessa afirmativa é, obviamente, o afastamento das operações de tratamento de informações sobre pessoas jurídicas (*v. g.*, sociedades empresárias, associações, entidades públicas) e entes não personalizados (*v. g.*, condomínio, massa falida, sociedade de fato) do regime de proteção de dados. Ainda que se levante hipóteses que estariam situadas numa zona gris entre dado referente a pessoa natural e pessoa jurídica, como a do nome empresarial derivado do nome de pessoa humana, ou o do endereço eletrônico corporativo em cuja constituição há nome do empregado,[98] há que se concluir pela não configuração de informação pessoal em tais casos, por consistir em dados referentes a pessoas jurídicas.[99]

Surge ainda como questão derivada merecedora da atenção do intérprete, o saber se dados de pessoa falecida são abarcados pela noção de dado pessoal, colhidos, assim, pelo regime de proteção de dados. O tema tem despertado especial interesse no bojo das discussões doutrinárias e jurisprudenciais sobre herança digital ou testamento digital.[100]

Se no contexto europeu não há muito o que se debater a respeito, haja vista o disposto nos Considerandos n. 27, 158 e 160 do GDPR, no direito brasileiro a falta de expressa previsão legal sobre o assunto instila o debate. Contudo, a interpretação sistemática dá firme sustentação à tese de que informações sobre pessoa falecida não possuem caráter pessoal para fins de aplicação do regime da LGPD. Dois são os principais argumentos: (i) com a morte da pessoa natural há o fim da personalidade jurídica,[101] de modo que

---

[98] A questão foi levantada em: ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 23-24.

[99] Depois de certa jurisprudência vacilante do TJUE, esta é a atual direção tomada no direito europeu. O Considerando n. 14 do GPDR não dá margem a dúvida: "*The protection afforded by this Regulation should apply to natural persons, whatever their nationality or place of residence, in relation to the processing of their personal data. This Regulation does not cover the processing of personal data which concerns legal persons and in particular undertakings established as legal persons, including the name and the form of the legal person and the contact details of the legal person*" (cf. PINHEIRO, Alexandre Sousa et. al. *Comentário ao Regulamento Geral de Proteção de Dados*. Coimbra: Almedina, 2018, p. 123-124). Em sentido diverso: cf. CORDEIRO, António Barreto Menezes. Dados pessoais: conceito, extensão e limites. *Revista de Direito Civil*, Coimbra, vol. 3, n. 2, p. 297-321, 2018, p. 307. Este autor, ademais, parece defender a compatibilidade da aplicação do regime de proteção de dados pessoais às pessoas coletivas no direito português, amparando-se no reconhecimento da tutela de direitos da personalidade desses entes: "O não reconhecimento de uma proteção aos dados pessoais de pessoas coletivas contrasta com os avanços recentes no campo dos direitos de personalidade clássicos. É hoje aceite, pacificamente, pelos tribunais portugueses que também as pessoas coletivas são titulares de alguns direitos de personalidade, caso do direito ao nome, o direito ao bom nome comercial e o direito à privacidade empresarial, que abrange, pelo menos, o sigilo da correspondência as particularidades de organização e de funcionamento e o *know-how* de uma entidade coletiva" (CORDEIRO, António Barreto Menezes. *Op. cit.*, p. 306-307).

[100] Sobre o assunto, v. MENDES, Laura Schertel; FRITZ, Karina Nunes. *Case Report*: Corte Alemã Reconhece a Transmissibilidade da Herança Digital. *Direito Público*, Porto Alegre, vol. 15, n. 85, p. 188–211, 2019; LEAL, Lívia Teixeira. *Internet e morte do usuário:* propostas para o tratamento *post mortem* do conteúdo inserido na rede. Rio de Janeiro: LMJ Mundo Jurídico, 2018; RESTA, Giorgio. La "morte" digitale. *Il Diritto dell'Informazione e dell'Informatica*, Milão, vol. XXIX, n. 6, p. 891–920, 2014.

[101] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 22. Sobre personalidade jurídica: MIRANDA, Pontes de. *Tratado de direito privado*. Rio de Janeiro: Borsoi, 1954, t. 1, p. 153-154; ANDRADE, Manuel A. Domingues de. *Teoria geral da relação jurídica*. reimp. Coimbra: Almedina, 1992, vol. 1, p. 30.

eventual tutela *post mortem* da privacidade, imagem e outras situações jurídicas existenciais fica a cargo do sistemática jurídico-civil aplicável; (ii) incompatibilidade com a teleologia da LGPD de tutela de direitos e liberdades fundamentais e o livre desenvolvimento da personalidade. Como já ressaltado em outra sede: "[a] pessoa falecida não corre [...] risco de discriminação, nem tampouco de ter o seu livre desenvolvimento prejudicado e, portanto, não faria sentido submeter seus dados ao mesmo sistema de proteção forte e preventivo estabelecido pela LGPD para as pessoas vivas".[102] Essa linha de pensamento, aliás, foi expressamente adotada pela Autoridade Nacional de Proteção de Dados (ANPD) na Nota Técnica nº 3/2023/CGF/ANPD.[103]

4. Dado anonimizado e dado pseudonimizado: contornos e regimes aplicáveis

Na interpretação construtiva da noção jurídica de dado pessoal, é importante se ter em mente que, por consequência, estabelece-se a linha divisória do que é e *não é* informação pessoal, ou seja, entre dado pessoal e dado não-pessoal. Se os dados não são relativos a pessoa natural identificada ou identificável, desde a origem ou após ulterior tratamento, são dados ditos *anônimos* ou que foram *anonimizados*, respectivamente.[104] A título exemplificativo, anônimos são os dados sobre substâncias poluentes coletados por sensores instalados em dispositivos de redes inteligentes (*smart grids*) para predição e monitoramento da qualidade do ar.[105] Anonimizados são, por sua vez, aqueles dados originalmente pessoais que foram objeto de tratamento – com técnicas e padrões vários – de maneira tal a retirar-lhes os identificadores diretos e indiretos. Com isso, os dados perdem, a princípio, o caráter pessoal. Tome-se o exemplo dos dados de geolocalização de usuários de telefone celular que, após submetidos a técnicas de anonimização de

[102] MENDES, Laura Schertel; FRITZ, Karina Nunes. *Case Report*: Corte Alemã Reconhece a Transmissibilidade da Herança Digital. *Direito Público*, Porto Alegre, vol. 15, n. 85, p. 188–211, 2019, p. 208.
[103] AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. *Nota Técnica nº 3/2023/CGF/ANPD*. Disponível em: gov.br. Em consulta feita pela Polícia Rodoviária Federal à ANPD sobre possibilidade de se criar um memorial em sítio eletrônico da corporação policial, com o fim homenagear os servidores já falecidos, a autoridade se manifestou pela não aplicação da normativa de proteção de dados pessoais a dados relativos a pessoas naturais já falecidas.
[104] Nos termos do artigo 5º, III, da LGPD, dado anonimizado é o "dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento".
[105] Para um estudo de caso do que ocorreu na cidade de Melbourne, Austrália, cf. SCHÜRHOLZ, Daniel; KUBLER, Sylvain; ZASLAVSKY, Arkady. Artificial intelligence-enabled context-aware air quality prediction for smart cities. *Journal of Cleaner Production*, vol. 271, 2020, p. 10-13.

dados, mais especificamente de agregação,[106] foram usados para a criação de mapas de calor e índice de isolamento social durante a pandemia de COVID-19 no Brasil.[107]

Se, de um lado, a qualificação do dado objeto de tratamento como *pessoal* tradicionalmente abre a porta de acesso ao regime de direitos e garantias ao titular dos dados, de outro, o dado qualificado como *não-pessoal* fica fora do âmbito material de aplicação de leis gerais de proteção de dados como a LGPD, o que empeça a aquisição de direitos do titular por indivíduos bem como a atribuição de situações jurídicas passivas (dever, encargos, sujeição etc.) do regime de proteção de dados a agentes de tratamento. A já referida parábola kafkiana é, neste ponto, igualmente ilustrativa do raciocínio: visto que o tratamento de dados não-pessoais não passa pelos umbrais da porta da lei (de proteção de dados), deveres, obrigações e encargos de controladores e operadores restam inaplicáveis.[108]

No que diz respeito aos dados anonimizados, em conformidade com o texto da LGPD, os dados não podem ter associação com pessoa identificada ou identificável de forma permanente e irreversível.[109] Se o "processo de anonimização ao qual [os dados pessoais] foram submetidos for revertido", seria, então, impróprio falar-se em dados não-pessoais. Nas últimas duas décadas, contudo, a própria ideia de duradoura irreversibilidade da

---

[106] Define-se agregação de dados, de acordo com o extinto Grupo de Trabalho do Artigo 29, da seguinte maneira: "*Aggregation and K-anonymity techniques aim to prevent a data subject from being singled out by grouping them with, at least, k other individuals. To achieve this, the attribute values are generalized to an extent such that each individual shares the same value. For example, by lowering the granularity of a location from a city to a country a higher number of data subjects are included. Individual dates of birth can be generalized into a range of dates or grouped by month or year. Other numerical attributes (e.g. salaries, weight, height, or the dose of a medicine) can be generalized by interval values (e.g. salary €20,000 – €30,000). These methods may be used when the correlation of punctual values of attributes may create quasi-identifiers*" (ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 5/2014 on Anonymisation techniques*. Bruxelas: [s. n.], 2014, p. 16).

[107] CONVERGÊNCIA DIGITAL. *Dispara número de estados e municípios que usam dados celulares na Covid-19*. Disponível em: convergenciadigital.com.br/. O tratamento de dados anonimizados com técnicas de agregação tem ampla incidência em aplicações tecnológicas para combater a pandemia de COVID-19 noutras partes do globo: cf. POOM, Age et al. COVID-19 is spatial: Ensuring that mobile Big Data is used for social good. *Big Data and Society*, vol. 7, n. 2, p. 1–7, jul.-dez./2020, p. 3; BUDD, Jobie et al. Digital technologies in the public-health response to COVID-19. *Nature Medicine*, vol. 26, n. 8, p. 1183–1192, ago./2020, p. 1184-1187.

[108] O Considerando n. 26 do GDPR expressamente aborda esse assunto: "[...] 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" (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: eur-lex.europa.eu/).

[109] A LGPD prescreve no *caput* do art. 12: "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", Vide também: MENDES, Laura Schertel. *Privacidade, proteção de dados e defesa do consumidor:* linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014, p. 57-58.

anonimização enfrenta contundentes críticas. Importantes estudos realizados no campo da ciência da computação, tais como as pesquisas de Latanya Sweeney,[110] Arvind Narayanan e Vitaly Shmatikov,[111] Gilbert Wondracek e outros,[112] Yves De Montjoye e outros,[113] demonstraram a existência de sérias limitações em práticas de anonimização de dados anteriormente reputadas confiáveis, principalmente num contexto de incremento das capacidades tecnológicas da análise de dados (*data analytics*). Isso levou a doutrina especializada a romper com o que Paul Ohm chamou de a "suposição da anonimização robusta"[114] (*robust anonymisation assumption*), que vigorava na teoria e prática da proteção de dados – a ideia de que com simples operações de eliminação ou substituição de atributos dos titulares dos dados respeitar-se-ia a privacidade, ao mesmo tempo em que seria reservada a utilidade das informações ao responsável pela base de dados.

Considerados os resultados dos estudos científicos e as pertinentes análises críticas, hoje há o reconhecimento de que sempre haverá fatores de risco de reidentificação de pessoas com o tratamento de dados anonimizados,[115] tendo em vista o enorme volume de dados disponibilizados publicamente (via internet) e o desenvolvimento da capacidade de processamento e análise de algoritmos de reidentificação.[116] No exame da robustez e do nível de garantia oferecidos por técnicas e práticas de anonimização de dados, sugere-se

---

[110] SWEENEY, Latanya. Simple Demographics Often Identify People Uniquely. *Carnegie Mellon University*, Data Privacy Working Paper 3, Pittsburgh, 2000.
[111] NARAYANAN, Arvind; SHMATIKOV, Vitaly. *How to break anonymity of the Netflix Prize dataset*. 2007, p. 3. Disponível em: citeseerx.ist.psu.edu/.
[112] WONDRACEK, Gilbert et al. A practical attack to de-anonymize social network users. *Proceedings – IEEE Symposium on Security and Privacy*, p. 223–238, 2010.
[113] DE MONTJOYE, Yves A. et al. Unique in the Crowd: The privacy bounds of human mobility. *Scientific Reports*, vol. 3, p. 1–5, 2013. De Montjoye publicou em coautoria outro trabalho em que analisa ataques de reidentificação a partir do uso de modelos generativos de aprendizado de máquina: ROCHER, L.; HENDRICKX, J. M.; DE MONTJOYE, Y. A. Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications*, vol. 10, n. 1, 2019.
[114] OHM, Paul. Broken promises of privacy: responding to the surprising failure of anonymization. *UCLA Law Review*, n. 57, p. 1701-1777, 2010, p. 1706.
[115] De acordo com a análise feita pelo Grupo de Trabalho de Proteção de Dados do Artigo 29, no Parecer nº 05/2014 sobre a anonimização de dados no contexto do direito da União Europeia, "à anonimização é inerente a existência de um fator de risco" (ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 5/2014 on Anonymisation techniques.* Bruxelas: [s. n.], 2014, p. 6-7).
[116] Quanto ao desenvolvimento dos algoritmos de reidentificação, Narayanan e Shmatikov asseveram: "*Re-identification algorithms are agnostic to the semantics of the data elements. It turns out there is a wide spectrum of human characteristics that enable re-identification: consumption preferences, commercial transactions, Web browsing, search histories, and so forth. Their two key properties are that (1) they are reasonably stable across time and contexts, and (2) the corresponding data attributes are sufficiently numerous and fine-grained that no two people are similar, except with a small probability*" (NARAYANAN, Arvind.; SHMATIKOV, Vitaly. Myths and fallacies of personally identifiable information. *Communications of the ACM*, vol. 53, n. 6, p. 24–26, 2010, p. 26). Os referidos autores, em recente reflexão sobre o atual estado da arte da reidentificação e seus impactos no campo das políticas legislativa e regulatória em matéria de proteção de dados, afirmam: "*Today's privacy regulations, including the GDPR, continue to put substantial weight on deidentification. Our key recommendation is that the burden of proof be on the data controller to affirmatively show that anonymized data cannot be linked to individuals, rather than on privacy advocates to show that linkage is possible*" (NARAYANAN, Arvind.; SHMATIKOV, Vitaly. Robust de-anonymization of large sparse datasets: a decade later. Disponível em: cs.princeton.edu).

que três tipos de riscos principais sejam levados em consideração:[117] distinção (*singling out*), possibilidade de ligação e inferência. O primeiro versa sobre a possibilidade de se isolar alguns ou todos os registros que destaca uma pessoa em uma base de dados; o segundo é a capacidade de se estabelecer uma conexão entre pelo menos dois registros relativos ao mesmo indivíduo ou mesmo grupo de pessoas; e o terceiro, por fim, diz com a possibilidade de inferir, com uma significativa probabilidade, o valor de um atributo a partir dos valores de um conjunto de outros atributos.[118] Entre as técnicas de anonimização utilizadas na atualidade, são já são bastante conhecidas, por exemplo, k-anonimato, l-diversidade e adição de perturbação.[119]

Bem pensadas as coisas, o intérprete deve considerar o estado da arte quanto ao processo de anonimização de dados e técnicas de reidentificação disponíveis, além da constatação da impossibilidade de cenário com risco zero.[120] Nessa direção deve ser compreendido o *critério dos esforços razoáveis* – ou "dos meios suscetíveis de ser razoavelmente utilizados", seu equivalente adotado no direito europeu – para reverter o(s) procedimento(s) de anonimização empregado(s), à luz do que dispõe o art. 12, *caput*, da LGPD, que, ademais, enseja certo elemento de abertura à *perspectiva relativa* do conceito de dado pessoal ao também dispor sobre a reversão "utilizando exclusivamente meios próprios" do agente de tratamento.[121] À semelhança do que fez o direito da UE ao

[117] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Op. cit.*, p. 11-12; 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: ipc.on.ca/.

[118] 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*, vol. 2, n. 4, 2020, p. 1.

[119] SMITH, Mick Smith; AGRAWAL, Rajeev. Anonymization Techniques. In: SCHINTLER, Laurie A.; MCNEELY, Connie L. (Orgs.). *Encyclopedia of Big Data*. Cham: Springer, 2022, p. 30-33.

[120] Cf. CARVALHO, Sérgio M.; FIORINI, Carolina. Dados não pessoais: a retórica da anonimização no enfrentamento à COVID-19 e o *privacywashing*. *Internet & Sociedade*, vol. 1, n. 2, p. 126–149, dez./2020, p. 135-139. Fink e Pallas argumentam que o parecer do antigo Grupo de Trabalho do Artigo 29 sobre técnicas de anonimização incorre em contradição, pois ao mesmo tempo que reconhece a adoção pelo GDPR – à época em debate no Parlamento europeu – de abordagem baseada no risco, parece propor um "teste de risco zero", o que, na prática, significa uma rejeição a tal tipo abordagem regulatória. FINCK, Michèle; PALLAS, Frank. They Who Must Not Be Identified – Distinguishing Personal from Non-Personal Data Under the GDPR. *International Data Privacy Law*, vol. 10, n. 1, p. 11–36, 2020, p. 15.

[121] Note-se que as teorias ou perspectivas objetivas (ou absoluta) e relativa são mais bem compreendidas como modelos ou esquemas teóricos ideais em relação aos quais a prática jurídica de proteção de dados pessoais (dimensões institucional-regulatória e jurisprudencial) pode se aproximar ou afastar, incorporando elementos de um e outro.

encampar esse critério[122] para avaliação da identificabilidade de dados, deve o direito brasileiro assegurar *análise contextual* dos elementos e fatores de risco de (re)identificabilidade (*v. g*., estado da arte das tecnologias computacionais, técnicas de reidentificação; custos, tempo e expertise necessárias para reverter a anonimização),[123] e fomentar a criação de padrões ou *standards*, notadamente a partir da atuação regulatória e diretiva da Autoridade Nacional de Proteção de Dados, para, ancorada em suas importantíssimas competências interpretativas da lei e de zelo pela proteção de dados pessoais (LGPD, art. 55-J, I, XX, e parágrafo único), balizar os contornos do que se deve considerar dados não vinculados a pessoa identificada ou identificável de forma "permanente e irreversível".

Em meio a essa discussão, com o advento das abordagens orientadas pela análise do risco, há o surgimento de propostas que visualizam entre o dado pessoal e o dado não-pessoal um gradiente de cores ou um *continuum* com categorias que se propõem a superar (em alguma medida) a lógica binária[124] dado pessoal/dado não-pessoal, informações a que se aplicam o regime de proteção de dados pessoais/informações a que não se aplicam o regime de proteção de dados pessoais. É nesse contexto que se coloca a ideia de *dado pseudonimizado*.

Segundo o Grupo de Trabalho de Proteção de Dados do Artigo 29, o procedimento de "pseudonimização consiste em substituir um atributo (tipicamente um atributo único) em um registro por outro";[125] seria um processo de mascaramento ou disfarce (*disguising*) de identidade,[126] que afeta principalmente identificadores diretos.[127] Nesse

---

[122] Entende-se que o referido "*critério*" dos esforços razoáveis, chamado por alguns no cenário europeu de "*teste*" (*v. g*,, ICO UK; Michèle Finck e Frank Pallas), deve ser compreendido como *conceito jurídico indeterminado normativo*, isto é, um conceito em larga medida incerto no seu conteúdo e extensão, dependente de preenchimento valorativo pelo aplicador do direito. ENGISCH, Karl. *Introdução ao pensamento jurídico*. 8. Ed. Trad. João Baptista Machado. Lisboa: Fundação Calouste Gulbenkian, 2001, p. 210-213.

[123] Ver LGPD, art. 12, § 1º.

[124] Nessa direção: SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review*, vol. 86, p. 1814–189, dez./2011, p. 1877; ESAYAS, Samson Yoseph. The role of anonymisation and pseudonymisation under the EU data privacy rules: beyond the 'all or nothing' approach. *European Journal of Law and Technology*, vol. 6, n. 2, 2015; POLONETSKY, Jules et. al. *The seven states of data: when is pseudonymous data not personal information?*. The Future of Privacy Forum, 2013. Disponível em fpf.org/.

[125] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 5/2014 on Anonymisation techniques*. Bruxelas: [s. n.], 2014, p. 20.

[126] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007, p. 18.

[127] ESAYAS, Samson Yoseph. *Op. cit.*, p. 4. Na mesma direção: *"Pseudonymization means that a true identifier such as name or patient identification number is replaced by a pseudonym that is unique to the individual but bears no relation to the person 'in the real world'. Pseudonym cannot therefore be used as a means of identification. This is because in pseudonymization, the information that reveals who the pseudonym relates to will be held securely, and separately, from the data being processed*" (NOUMEIR, Rita; LEMAY, Alain; LINA, Jean-Marc. Pseudonymization of Radiology Data for Research Purposes. *Journal of Digital Imaging*, vol. 20, n. 3, p. 284-295, 2007, p. 286).

sentido, a pseudonimização opera de maneira que as informações não podem ser conectadas a um titular de dados específico sem que se recorra a informações suplementares, desde que estas sejam mantidas separadamente, empregadas medidas administrativas, organizacionais e de segurança. Em pesquisas médicas, se estuda a implementação dessas técnicas a fim de se alcançar o respeito ao direito à proteção de dados pessoais dos sujeitos da pesquisa e o tratamento de dados de modo útil à investigação científica. Para estudos de radiologia clínica, a título de exemplo, a pseudonimização é importante para identificar dados de certo paciente de forma consistente com o passar do tempo – o que a anonimização ainda não faz.[128] O pseudônimo de um paciente específico é associado a todas as informações de-identificadas relativas a ele apesar do momento em que houve a de-identificação dos dados.[129] Como resultado há o acompanhamento do evolver do quadro clínico do sujeito da pesquisa.[130]

Essa maneira de conceber a pseudonimização como "substituição de identidade" (*ID replacement*)[131] parece ter encontrado acolhida tanto no texto da LGPD (art. 13, § 2º[132]), como no regulamento geral europeu (GDPR, art. 4º, 5[133]). Conquanto haja autores e autoridades governamentais que defendam que a pseudonimização seja mais uma técnica de anonimização,[134] a distinção entre ambos os processos é o que parece prevalecer

[128] NOUMEIR, Rita; LEMAY, Alain; LINA, Jean-Marc. *Op. cit.*, p. 286-287. De forma didática, explica Jaap-Henk Hopeman: *"Ideally, pseudonyms are unique – meaning that no two people share the same pseudonym in a given context. This is, for example, the case for email addresses and account names, but not necessarily the case for freely chosen nicknames. The fact that pseudonyms are unique allows them to be used to single out a person: every time a system observes a certain pseudonym, it knows this pseudonym belongs to the same person, even though it may not be able to tell who that person is (yet). Pseudonyms allow us to link events or data that belong to the same person (without knowing their true identity)*" (HOEPMAN, Jaap-Henk. *Privacy is hard and seven other myths:* achieving privacy through careful design. Cambridge: The MIT Press, 2021, p. 21).

[129] NOUMEIR, Rita; LEMAY, Alain; LINA, Jean-Marc. *Op. cit.*, p. 287.

[130] Para mais exemplos e casos de aplicação da pseudonimização no setor da saúde, cf. EUROPEAN AGENCY FOR CYBERSECUIRTY. *Deploying pseudonymisation techniques:* the case of the health sector. [S.l.]: ENISA, 2022, p. 12 et. seq.

[131] 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*, vol. 10, n. 1, p. 11–36, 2020, p. 22.

[132] Textualmente se lê: "Para os efeitos deste artigo, a pseudonimização é o tratamento por meio do qual um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo, senão pelo uso de informação adicional mantida separadamente pelo controlador em ambiente controlado e seguro".

[133] À luz do GDPR pseudonimização é "o tratamento de dados pessoais de forma que deixem de poder ser atribuídos a um titular de dados específico sem recorrer a informações suplementares, desde que essas informações suplementares sejam mantidas separadamente e sujeitas a medidas técnicas e organizativas para assegurar que os dados pessoais não possam ser atribuídos a uma pessoa singular identificada ou identificável".

[134] ESAYAS, Samson Yoseph. The role of anonymisation and pseudonymisation under the EU data privacy rules: beyond the 'all or nothing' approach. *European Journal of Law and Technology*, vol. 6, n. 2, 2015, p. 4; GARFINKEL, Simson L. *De-Identification of Personal Information*. [S.l.]: National Institute of Standards and Technology, 2015, p. 2; PERSONAL DATA PROTECTION COMMISSION. *Guide to basic anonymisation*. Singapura: PDPC Singapore, 2022. Disponível em: pdpc.gov.sg/.

em sede legal, tendo em consideração a LGPD[135] e, mais explicitamente, o GDPR.[136]

No mais das vezes, não se discute que o dado pseudonimizado consiste em dado de caráter pessoal,[137] é dizer, relativo a pessoa identificada ou identificável, à luz do conceito amplo.[138] Uma vez configurado tratamento de dado pessoal, a implementação de técnicas de pseudonimização (*v. g.,* uso de função *hash* para cifrar identificadores diretos numa base de dados[139]), na verdade, se traduz em medidas de mitigação de risco envolvendo a atividade de tratamento em questão. Tome-se o simples exemplo da cifragem com criptografia forte de certa base de dados pelo controlador: em caso de incidente de segurança por vazamento de dados, reduz-se significativamente o risco de lesão a direitos e liberdades fundamentais dos titulares. Tal abordagem baseada em avaliação de riscos aproxima, ainda, a pseudonimização de dados do princípio da proteção de dados desde a concepção, ao ser um processo que pode ser incorporado desde o desenvolvimento de produtos e serviços (LGPD, art. 46, § 2º).[140]

Por fim, cumpre destacar que, em vista da diversidade de cenários de uso de técnicas de pseudonimização e complexidade do tema, além da escassa disciplina da LGPD a respeito, há a necessidade de aprofundamento do estudo e análise crítica no direito brasileiro quanto ao regime jurídico por vezes modulado ou particularizado aplicável a dados pseudonimizados. Considere-se, por exemplo, situações em que os dados suplementares que permitem a identificação das informações pelo controlador não sejam mais mantidos por ele ou por algum operador, ou tenham sido descartados. Inalterada a natureza pessoal de tais dados pseudonimizados, será possível o exercício de direitos como direito de acesso e de retificação, perante o controlador? Seria correto dizer que, a

---

[135] Vide artigos 5º, III e XI, e 13, §4º.

[136] Vide Considerandos n. 26 a 29, 75, 78; e artigo 4º, 5. Na mesma linha: EUROPEAN AGENCY FOR CYBERSECUIRTY. *Deploying pseudonymisation techniques:* the case of the health sector. [S.l.]: ENISA, 2022, p. 9; INFORMATION COMMISSIONER'S OFFICE. *DRAFT anonymisation, pseudonymisation and privacy enhancing technologies guidance*. Wilmslow: ICO, 2022. Disponível em: ico.org.uk/; FINCK, Michèle; PALLAS, Frank. *Op. cit.*, p. 21.

[137] Tem-se ciência de que há quem discuta a possibilidade de se considerar dado pseudonimizado como dada anonimizado: cf. MOURBY, M. et al. Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK. *Computer Law and Security Review*, vol. 34, n. 2, p. 222–233, 2018.

[138] Na doutrina brasileira: cf. BIONI, Bruno R. Compreendendo o conceito de anonimização e dado anônimo. *Revista do Advogado,* São Paulo, vol. 39, n. 144, p. 22-32, nov./2019, p. 26; 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. 180; FRAZÃO, Ana; CARVALHO, Angelo Prata de; MILANEZ, Giovanna. *Curso de proteção de dados pessoais:* fundamentos da LGPD. Rio de Janeiro: Forense, 2022, p. 62.

[139] Por todos: AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. *Introducción al* hash *como técnica de seudonimización de datos personales*. [S.l.]: AEPD, 2019. Disponível em: aepd.es/.

[140] O GDPR reconhece tais características da pseudonimização, medida de gerenciamento de riscos e mecanismo de promoção da proteção de dados desde a concepção (*data protection by design*), de forma mais evidente, por exemplo, em seus Considerandos n. 28 e 78, e arts. 25 e 35. Nessa direção: FINCK, Michèle; PALLAS, Frank. They Who Must Not Be Identified – Distinguishing Personal from Non-Personal Data Under the GDPR. *International Data Privacy Law*, vol. 10, n. 1, p. 11–36, 2020, p. 21.

depender das circunstâncias da pseudonimização, o agente de tratamento pode ser dispensado de observar certas disposições sobre direitos dos titulares?

5. Considerações finais

Após análise do disputado conceito de dado pessoal no cenário jurídico internacional e brasileiro, tomadas contribuições e aportes especialmente da experiência jurídica da UE, pode-se compreender alguns de seus contornos e apontar, fundamentadamente, algumas direções para a interpretação e aplicação das normas que lhe são atinentes no texto da LGPD e no sistema pátrio. A perspectiva expansionista adotada na normativa brasileira se expressa no conceito amplo de informação pessoal que, não obstante a simplicidade e concisão dos seus termos, enseja não poucas complexidades das abordagens objetiva e relativa no que tange à identificabilidade do titular dos dados, notadamente a partir do critério dos meios suscetíveis de ser razoavelmente utilizados, isto é, dos esforços razoáveis, para a (re)identificação do titular dos dados.

Como o título deste trabalho indica e a investigação empreendida denota, não se está a delinear enunciados hermenêuticos conclusivos e exaurientes a respeito da noção de dado pessoal e suas repercussões jurídicas. As anotações nesta sede formuladas visaram traçar (certos) confins e balizas no afã de contribuir para importante discussão de que a doutrina nacional deve se ocupar e auxiliar às instituições incumbidas da interpretar e aplicar de maneira sistemática a normativa de proteção de dados pessoais no Brasil, notadamente a ANPD, em suas competências de interpretação da LGPD e de zelo pela proteção de dados pessoais, e também o Poder Judiciário, em seu mister constitucional de tutela jurisdicional de direitos.

Referências

ADRIAANS, Pieter. Information. *The Stanford Encyclopedia of Philosophy* – Edward N. Zalta (ed.). Disponível em: plato.stanford.edu/.

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. *Introducción al* hash *como técnica de seudonimización de datos personales*. [S.l.]: AEPD, 2019. Disponível em: aepd.es/.

ALBERS, Marion. A complexidade da proteção de dados. *Direitos Fundamentais & Justiça*, vol. 10, n. 35, 2016.

A-LEVELS and GCSEs: How did the exam algorithm work? *BBC,* [s.l.], 20 ago. 2020. Disponível em: bbc.com/.

ANDRADE, Manuel A. Domingues de. *Teoria geral da relação jurídica*, vol. 1. reimp. Coimbra: Almedina, 1992.

ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 4/2007 on the concept of personal data*. Bruxelas: [s. n.], 2007. Disponível em: ec.europa.eu/.

ARTICLE 29 DATA PROTECTION WORKING PARTY. *Privacy on the Internet – An integrated EU Approach to On-line Data Protection*. Bruxelas: [s. n.], 2000. Disponível em: ec.europa.eu/.

ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 2/2010 on online behavioral advertising*. Bruxelas: [s. n.], 2010. Disponível em: ec.europa.eu/.

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. *Nota Técnica nº 3/2023/CGF/ANPD*. Disponível em: gov.br/.

BENNETT, Colin J.; RAAB, Charles D. *The Governance of Privacy*: Policy Instruments in Global Perspective. 1. Ed. London-New York: Routledge, 2003.

BIONI, Bruno. *Xeque-Mate:* o tripé de proteção de dados pessoais no xadrez das iniciativas legislativas no Brasil. São Paulo: GPOPAI, 2015.

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

BIONI, Bruno R. Compreendendo o conceito de anonimização e dado anônimo. *Revista do Advogado,* São Paulo, vol. 39, n. 144, nov./2019.

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*, vol. 32, n. 2, 2016.

BORGESIUS, Frederik Zuiderveen. The Breyer Case of the Court of Justice of the European Union: IP Addresses and the Personal Data Definition. *European Data Protection Law Review*, vol. 3, n. 1, 2017.

BENNETT, Colin J.; RAAB, Charles D. *The Governance of Privacy*: Policy Instruments in Global Perspective. 1. ed. London-New York: Routledge, 2003.

BUTLER, Brian E. Legal Pragmatism: Banal or Beneficial as a Jurisprudential Position? *Essays in Philosophy*, vol. 3, n. 2, jun./2002.

CARVALHO, Sérgio M.; FIORINI, Carolina. Dados não pessoais: a retórica da anonimização no enfrentamento à COVID-19 e o *privacywashing*. *Internet & Sociedade*, vol. 1, n. 2, dez./2020.

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: ipc.on.ca/.

CORDEIRO, António Barreto Menezes. Dados pessoais: conceito, extensão e limites. *Revista de Direito Civil*, Coimbra, vol. 3, n. 2, 2018.

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: search.coe.int/.

CRAWFORD, Kate; BOYD, dana. Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, vol. 15, n. 5, 2012.

CUEVA, Ricardo Villas Bôas. Proteção de dados pessoais no Judiciário. *Revista do Advogado*, São Paulo, n. 144, nov./2019.

CUKIER, Kenneth; MAYER-SCHÖNBERGER, Viktor. *Big Data:* A Revolution That Will Transform How We Live, Work, and Think. Boston: Houghton Mifflin Harcourt, 2013.

DE MONTJOYE, Yves-Alexandre et al. Unique in the Crowd: The privacy bounds of human mobility. *Scientific Reports*, vol. 3, 2013.

DONEDA, Danilo. *Da privacidade à proteção de dados pessoais*. 2. ed. São Paulo: Revista dos Tribunais, 2020.

DUHIGG, Charles. How Companies Learn Your Secrets. *The New York Times Magazine,* 16 fev. 2012. Disponível em: nytimes.com/.

DWORKIN, Ronald. *O império do direito*. Trad. Jefferson Luiz Camargo. São Paulo: Martins Fontes, 1999.

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

ESTADOS UNIDOS. *Children's Online Privacy Protection Act of 1998*. Disponível em: law.cornell.edu/.

EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS. *Handbook on European Data Protection Law*. Luxembourg: Publications Office of the European Union, 2018.

EUROPEAN AGENCY FOR CYBERSECUIRTY. *Deploying pseudonymisation techniques:* the case of the health sector. [S.l.]: ENISA, 2022. Disponível em: enisa.europa.eu/.

FACIAL recognition: School ID checks lead to GDPR fine. *BBC News,* 27 ago. 2019. Disponível em: bbc.com/.

FINCK, Michèle; PALLAS, Frank. They Who Must Not Be Identified – Distinguishing Personal from Non-Personal Data Under the GDPR. *International Data Privacy Law*, vol. 10, n. 1, 2020.

FINOCCHIARO, Giusella. Anonimato. In: *Digesto delle Discipline Privatistiche – Sezione Civile*. Aggiornamento. Torino: UTET, 2010.

FLORIDI, Luciano. The ontological interpretation of informational privacy. *Ethics and Information Technology*, vol. 7, n. 4, 2005.

FRAZÃO, Ana; CARVALHO, Angelo Prata de; MILANEZ, Giovanna. *Curso de proteção de dados pessoais:* fundamentos da LGPD. Rio de Janeiro: Forense, 2022.

GARFINKEL, Simson L. *De-Identification of Personal Information*. [S.l.]: National Institute of Standards and Technology, 2015. Disponível em: nvlpubs.nist.gov/.

GITELMAN, Lisa (Org.). *'Raw data' is an oxymoron*. Cambridge; London: The MIT Press, 2013.

GRATTON, Éloïse. *Understanding Personal Information: Managing Privacy Risks*. Markham: LexisNexis, 2013.

GRAY, Stacey. A Closer Look at Location Data: Privacy and Pandemics. Future of Privacy Forum, 2020. Disponível em: fpf.org/.

HILDEBRANDT, Mireille. *Smart Technologies and The End(s) of Law*. Cheltenham-Northampton: Edward Elgar, 2015.

HOEPMAN, Jaap-Henk. *Privacy is hard and seven other myths:* achieving privacy through careful design. Cambridge: The MIT Press, 2021.

HOOFNAGLE, Chris J.; VAN DER SLOOT, Bart; BORGESIUS, Frederik Z. The European Union general data protection regulation: what it is and what it means. *Information & Communications Technology Law*, vol. 28, n. 1, 2019.

INTERNATIONAL COMMITTEE OF THE RED CROSS. *Handbook on Data Protection in Humanitarian Action*. 2. ed. Geneva: ICRC, 2020.

KERR, Ian. Foreword. In: GRATTON, Éloïse. *Understanding Personal Information: Managing Privacy Risks*. Markham: LexisNexis, 2013.

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

MACHADO, Diego C.; SOUZA, Carlos Affonso P. Tutela da privacidade, guarda de registros e portas lógicas no direito brasileiro. In: FERRARI, Isabela; BECKER, Daniel (Org.). *Regulação 4.0* – Novas tecnologias sob a perspectiva regulatória. São Paulo: Revista dos Tribunais, 2019.

MANTELERO, Alessandro. Responsabilità e rischio nel Reg. UE 2016/679. *Le Nuove Leggi Civili Commentate*, vol. XL, n. 1, 2017.

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. Indaiatuba: Editora Foco, 2022.

MACHADO, Diego; MENDES, Laura Schertel. A proteção dos dados sensíveis inferidos: um comentário ao Caso C-184/20 do Tribunal de Justiça Europeu. *Revista de Direito do Consumidor*, São Paulo, vol. 144, a. 31, nov.-dez./2022.

MENDES, Laura Schertel. *Privacidade, proteção de dados e defesa do consumidor:* linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014.

MENDES, Laura Schertel. *Habeas data* e autodeterminação informativa: os dois lados da mesma moeda. *Direitos Fundamentais & Justiça*, vol. 12, n. 39, 2018.

MENDES, Laura Schertel; FRITZ, Karina Nunes. *Case Report*: Corte Alemã Reconhece a Transmissibilidade da Herança Digital. *Direito Público*, Porto Alegre, vol. 15, n. 85, 2019.

MIRANDA, Pontes de. *Tratado de direito privado*. Rio de Janeiro: Borsoi, 1954, t. 1.

MOURBY, M. et al. Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK. *Computer Law and Security Review*, vol. 34, n. 2, 2018.

NARAYANAN, Arvind.; SHMATIKOV, Vitaly. Myths and fallacies of personally identifiable information. *Communications of the ACM*, vol. 53, n. 6, 2010.

NARAYANAN, Arvind.; SHMATIKOV, Vitaly. Robust de-anonymization of large sparse datasets: a decade later. Disponível em: cs.princeton.edu/.

OFQUAL. *Executive Summary – Awarding GCSE, AS, A level, advanced extension awards and extended project qualifications in summer 2020:* interim report. Londres: [s. n.], 13 ago. 2020. Disponível em: assets.publishing.service.gov.uk/.

OHM, Paul. Broken promises of privacy: responding to the surprising failure of anonymization. *UCLA Law Review*, n. 57, 2010.

OLEJNIK, Lukasz; ACAR, Gunes; CASTELLUCCIA, Claude; DIAZ, Claudia. *The leaking battery A privacy analysis of the HTML5 Battery Status API*. Disponível em: eprint.iacr.org/.

PEARCE, Henry. Big data and the reform of the European data protection framework: an overview of potential concerns associated with proposals for risk management-based approaches to the concept of personal data. *Information and Communications Technology Law*, vol. 26, n. 3, 2017.

PERSONAL DATA PROTECTION COMMISSION. *Guide to basic anonymisation*. Singapura: PDPC Singapore, 2022. Disponível em: pdpc.gov.sg/.

PINHEIRO, Alexandre Sousa et. al. *Comentário ao Regulamento Geral de Proteção de Dados*. Coimbra: Almedina, 2018.

PINTO, Paulo Mota. O direito ao livre desenvolvimento da personalidade. *Boletim da Faculdade de Direito da Universidade de Coimbra*. Número especial Portugal-Brasil ano 2000. Coimbra: Coimbra Editora, 1999.

POLONETSKY, Jules et. al. *The seven states of data: when is pseudonymous data not personal information?*. The Future of Privacy Forum, 2013. Disponível em fpf.org/.

POSNER, Richard A. *Overcoming Law*. Cambridge: Harvard University Press, 1995.

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*, vol. 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*, vol. 10, n. 1, 2018.

RESTA, Giorgio. La "morte" digitale. *Il Diritto dell'Informazione e dell'Informatica*, Milão, vol. XXIX, n. 6, 2014.

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

SMITH, Mick Smith; AGRAWAL, Rajeev. Anonymization Techniques. In: SCHINTLER, Laurie A.; MCNEELY, Connie L. (Orgs.). *Encyclopedia of Big Data*. Cham: Springer, 2022.

THOMPSON, Stuart A.; WARZEL, Charlie. The Privacy Project – Twelve Million Phones, One Dataset, Zero Privacy. *The New York Times*, 19 dez. 2019. Disponível em: nytimes.com/.

SCHÜRHOLZ, Daniel; KUBLER, Sylvain; Zaslavsky, Arkady. Artificial intelligence-enabled context-aware air quality prediction for smart cities. *Journal of Cleaner Production*, vol. 271, 2020.

SCHWARTZ, Paul M.; SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review*, vol. 86, dec./2011.

SOLOVE, Daniel J. *Understanding Privacy*. Cambridge; London: Harvard University Press, 2008.

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*, vol. 7, 2016.

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: op.europa.eu/.

UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia. Segunda Seção. *Case C-434/16*, *Peter Novak v. Data Protection Commissioner*. Luxemburgo, 20 dez. 2017.

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.

WACKS, Raymond. *Personal Information:* Privacy and the Law. Oxford: Oxford University Press, 1989.

WACHTER, Sandra; MITTELSTADT, Brent D. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, vol. 2019, n. 2, 2019.

ZENO-ZENCOVICH, Vincenzo. Informazione (profili civilistici). In: *Digesto* – Sezione Civile, vol. IX. Torino: UTET, 1993.

Como citar:

MACHADO, Diego. Considerações iniciais sobre o conceito de dado pessoal no ordenamento jurídico brasileiro. **Civilistica.com**. Rio de Janeiro, a. 12, n. 1, 2023. Disponível em: <http://civilistica.com/consideracoes-iniciais-sobre-o-conceito/>. Data de acesso.

civilistica.com

Recebido em:
1.8.2022
Aprovado em:
13.4.2023