

respaldada por um sistema ou uma cláusula geral de proteção. Essa orientação é energizada pela concepção de um direito privado despatrimonializado ou repersonalizado.

Por isso, necessário se faz revisitar constantemente os direitos da personalidade para se aperfeiçoar a busca incessante e mutável da tutela da pessoa humana⁵⁷. As novas tecnologias trazem novos desafios a esse respeito⁵⁸.

2.2 A PROJEÇÃO DA PERSONALIDADE POR MEIO DOS DADOS

2.2.1 Dados pessoais e projeção da personalidade: uma nova identidade

Personalidade significa as “características ou o conjunto de características que distingue uma pessoa”⁵⁹ da outra. Com base nessa abordagem semântica, os direitos da personalidade seriam os caracteres incorpóreos e corpóreos⁶⁰ que conformam a projeção⁶¹ da pessoa humana. Nome, honra, integridade física e psíquica⁶² seriam apenas alguns dentre uma série de outros atributos que dão forma a esse prolongamento⁶³.

⁵⁷ CAPELO DE SOUSA, Rabindranath V. A. Op. cit. p. 764.

⁵⁸ CORDEIRO, António Menezes. Op. cit., p. 61: “A dramática experiência do III Reich levou, na República Federal Alemã subsequente a 1949, a um surto no desenvolvimento de direitos fundamentais e de personalidade. Esse movimento intensificou-se com a multiplicação dos meios suscetíveis de agredir ou pôr em causa a esfera mais pessoal de cada um: meios de comunicação e de vigilância, informática e biotecnologia, como exemplos”.

⁵⁹ HOUAISS, Antônio; VILLAR, Mauro de Salles. Op. cit., p. 1.480.

⁶⁰ SIMÃO, José Fernando. *Responsabilidade civil do incapaz*. São Paulo: Atlas, 2008. p. 14.

⁶¹ BITTAR, Carlos Alberto. *Os direitos...* Op. cit., p. 1: “Consideram-se como da personalidade os direitos reconhecidos à pessoa humana tomada em si mesma e em suas projeções na sociedade, previstos no ordenamento jurídico exatamente para a defesa de valores inatos no homem, como a vida, a higidez física, a intimidade, a honra, a intelectualidade e outros tanto”. Em sentido análogo: “Nesta ordem de ideias, há que concluir que a personalidade jurídica é a projecção no Direito (no mundo normativo jurídico) da personalidade humana”. (CARVALHO, Orlando de. Op. cit., p. 190).

⁶² Esses são, por exemplo, algumas das espécies de direitos da personalidade listadas pelo CC.

⁶³ FRANÇA, Rubens Limongi. Direitos da personalidade. In: MENDES, Gilmar Ferreira; STOCO, Rui (Org.). *Doutrinas essenciais: direito civil – parte geral – pessoas e domicílio*. São Paulo: Revista dos Tribunais, 2011. p. 654: “Portanto, Direitos da Personalidade dizem-se as faculdades jurídicas cujo objeto são os diversos aspectos da própria pessoa do sujeito, bem assim seus prolongamentos e projeções. Ao tra-

Dada a *ipseidade* que difere o ser humano dos outros entes e entre seus próprios pares (*distinctum subsistens*)⁶⁴, a ciência jurídica o protege das agressões que afetem a sua individualidade⁶⁵. Trata-se de conferir tutela jurídica aos elementos que emprestam conteúdo ao valor-fonte do ordenamento jurídico, aos bens (da personalidade) que individualizam o sujeito perante a sociedade⁶⁶.

Sob essa perspectiva, um dado, atrelado à esfera de uma pessoa⁶⁷, pode se inserir dentre os direitos da personalidade. Para tanto, ele deve ser adjetivado

tarmos dos Direitos da Personalidade em espécie, será elucidada a parte do nosso conceito relativo às palavras 'prolongamentos' e 'projeções'.

⁶⁴ GONÇALVES, Diogo Costa. *Pessoa e direitos da personalidade: fundamentação ontológica*. Coimbra: Almedina, 2008. p. 44: "Mas, sendo *subsistens*, o Homem é *distinctum subsistens*. Quando afirmamos *distinctum*, porque o acto de ser, no Homem, está marcado pela autoporessão, pela imanência, pela ipseidade... em suma, por uma radical intimidade ontológica que só toda a densidade semântica do 'eu sou' é capaz de expressar. Nesta acepção, que evidencia a realidade da pessoa como ser em si podemos afirmar que o Homem é o único *subsistens* em sentido estrito ou próprio".

⁶⁵ TEPEDINO, 2008, p. 29: "De outro ponto de vista, todavia, tem-se personalidade como conjunto de características e atributos da pessoa humana, considera-se como objeto de proteção por parte do ordenamento jurídico. A pessoa, vista deste ângulo, há de ser tutelada das agressões que afetam a sua personalidade (...) Dito diversamente, considerada como sujeito de direito, a personalidade não pode ser dele o seu objeto. Considerada, ao revés, como valor, tendo em conta o conjunto de atributos inerentes e indispensáveis ao ser humano (que se irradiam da personalidade), constituem bens jurídicos em si mesmos, dignos de tutela privilegiada".

⁶⁶ CUPIS, Adriano de. *Os direitos da personalidade*. Tradução Afonso Celso Furtado Rezende. São Paulo: Quorum, 2008. p. 180: "O indivíduo como unidade da vida social e jurídica, tem necessidade de afirmar a própria individualidade, distinguindo-se dos outros indivíduos, e, por consequência, ser conhecido por quem é na realidade. O bem que satisfaz essa necessidade é o da identidade, o qual consiste, precisamente, no distinguir-se das outras pessoas nas relações sociais".

⁶⁷ MENDES, Laura Schertel. O direito fundamental à proteção de dados pessoais. *Revista de Direito do Consumidor*, ano 20, v. 79, p. 75, jul./set. 2011: "Tendo em vista tratar-se de direito à personalidade, já que os dados armazenados representam a pessoa na sociedade". Em sentido similar: CATALA, Pierre. *Ebauche d'une théorie juridique de l'information*, p. 20, *apud* DONEDA, Danilo. *Da privacidade...* Op. cit., p. 157: "Mesmo que a pessoa em questão não seja a 'autora' da informação, no sentido de sua concepção, ela é titular legítima dos seus elementos. Seu vínculo com o indivíduo é por demais estreito para que pudesse ser de outra forma. Quando o objeto dos dados é um sujeito de direito, a informação é um direito da personalidade".

como pessoal, caracterizando-se como uma projeção, extensão ou dimensão do seu titular⁶⁸.

E, nesse sentido, cada vez mais, as atividades de processamento de dados têm ingerência na vida das pessoas. Hoje vivemos em uma sociedade e uma economia que se orientam e movimentam a partir desses *signos identificadores* do cidadão.

Trata-se de um novo tipo⁶⁹ de identidade⁷⁰ e, por isso mesmo, tais *dossiês digitais*⁷¹ devem externar informações corretas para que seja fidedignamente *projetada* a identidade⁷² do titular daquelas informações.

Isso acaba por justificar *dogmaticamente* a inserção dos dados pessoais na categoria dos direitos da personalidade, assegurando, por exemplo, que uma pessoa exija a retificação de seus dados pessoais para que a sua projeção seja precisa⁷³.

⁶⁸ RODOTÀ, Stefano. *Il diritto di avere*. Roma: Laterza, 2012. p. 148: “Nel discorso giuridico, la registrazione di un dato deve essere compagna da una riflessione sul senso del passaggio nella nuova dimensione, dunque da una attività necessariamente ricostruttiva”.

⁶⁹ TENNIS, Bradley. Privacy and identity in a networked world. In: AKRIVOPOLOUS, Christina; PSYGKAS, Athanasios (Org.). *Personal data privacy and protection in a surveillance era: technologies and practices*. New York: Information Science Reference, 2011. p. 12-13.

⁷⁰ RODOTÀ, Stefano. *Il diritto...* Op. cit., p. 314: “Tutto questo, oggi, può essere considerato anche nella dimensione de diritti, di una costruzione dell’indetità che finisce com il coincidere con la costruzione stessa dell’umano (...) costruire liberamente la própria indetità utilizzando tutte le opportunità socialmente disponibili. La nuova dimensione dell’umano sige uma diversa misura giuridica, che dilata l’ambito de diritti fondamentali della persona”.

⁷¹ A expressão de Daniel J. Solove, cujo título da obra expressa a atribuição de uma nova identidade provinda dos dados pessoais. SOLOVE, Daniel. J. *The digital person: technology and privacy in the information age*. New York: New York University Press, 2004. p. 3.

⁷² Enquadrando direito à identidade como um dos direitos da personalidade, veja RIZZARDO, Arnaldo. *Parte geral do código civil*. Rio de Janeiro: Forense, 2011. p. 157: “O ponto de realce, aqui, restringe-se à identidade como direito da personalidade, para que tenham as pessoas uma presença na sociedade e perante o Estado, e para que não se considerem simples quantidades ou números na ordem do gênero humano. Por isso, decorre do direito da personalidade o tratamento como uma individualidade, dentro de uma performance própria – ou um ser que se destaca, merecedor de respeito e titular de uma posição única”.

⁷³ Como decorrência do princípio da qualidade dos dados, o cidadão tem o direito de exigir a correção de dados incompletos, inexatos ou desatualizados, de acordo com os arts. 6º, V, e 8º, III, ambos da LGPD.

Por isso, os dados pessoais não estão relacionados somente com a privacidade, transitando dentre mais de uma das espécies dos direitos da personalidade. Tal construção dogmática⁷⁴ é útil, pois é tal *ampliação normativa* que assegura o direito à retificação e de acesso aos dados e outras posições jurídicas próprias do direito à proteção dos dados pessoais (e.g., direito de revisão de decisões automatizadas)⁷⁵⁻⁷⁶.

Seria contraproducente e até mesmo incoerente pensar a proteção de dados pessoais somente sob as lentes do direito à privacidade. O eixo da privacidade está ligado ao controle de informações pessoais do que seja algo íntimo ou privado do sujeito. A proteção dos dados pessoais não se satisfaz com tal técnica normativa, uma vez que a informação pode estar sob a esfera pública, discutindo-se, apenas, a sua exatidão, por exemplo⁷⁷.

⁷⁴ Apesar de discordar com a conclusão final – direito à identidade informacional – concorda-se com o trajeto percorrido pelo jurista português que sublinha a importância do direito à proteção dos dados pessoais sob as lentes do direito da personalidade para fins de uma correta construção dogmática. Como será visto mais à frente, o reconhecimento do direito à proteção dos dados pessoais como um novo direito da personalidade já é suficiente para galgar tal rigor dogmático e, nesse sentido, o que propomos nesse trabalho é reavaliar a sua percepção tradicional – autodeterminação informacional – para uma recepção crítica de acordo com a cultura jurídica legal brasileira. PINHEIRO, Alexandre Sousa. *Privacy e proteção de dados pessoais: a construção dogmática do direito à identidade informacional*. Lisboa: AAFDL, 2015. p. 827-829.

⁷⁵ Toda vez que submetido a um processo de decisão totalmente automatizado, sem que haja qualquer tipo de intervenção humana, o cidadão tem assegurado o direito de solicitar sua revisão, bem como obter explicação a respeito dos critérios que estão sendo utilizados para tal automatização (art. 10, *caput* e § 1º, da LGPD).

⁷⁶ RODATÀ, Stefano. Persona, riservatezza, identità. *Rivista Critica del Diritto Privato*, ano XV, p. 583-584, dic. 1997: “Alcuni abituali usi linguistici, e il peso di una lunga vicenda d’origine, inducono a considerare la legge sulla protezione dei dati personali come una disciplina della ‘libertà informatica’ come un insieme di norma sulla ‘privacy informatica’. Ma è sicuramente improprio l’uso sia del sostantivo (privacy) (...) il trattamento dei dati personali si svolga nel rispetto dei diritti, della libertà fondamentali, nonché della dignità delle persone fisiche, con particolare riferimento alla riservatezza e all’identità personale (...) Una lettura condotta unicamente in termini di privacy non sarebbe soltanto riduttiva, ma porterebbe pure ad una impropria ricostruzione di questo stesso concetto(...) entrambe confermano la dilatazione della considerazione normativa”.

⁷⁷ ANDRADE, Norberto Nuno Gomes de Andrade. The right privacy and the right to identity in the Age of ubiquitous computing: Friends or foes? A proposal towards a legal articulation. In: AKRIVOPOLOUS, Christina; PSYGKAS, Athanasios (Org.). *Personal data privacy and protection in a surveillance era: technologies and practices*.

Ao lado do princípio da qualidade dos dados, o direito de correção é uma construção que deriva da perspectiva da identidade do sujeito e não do direito à privacidade⁷⁸. É o primeiro direito de personalidade que determina a necessidade de haver uma correspondência⁷⁹ fidedigna entre a pessoa e seus dados pessoais. A esfera do que é público ou privado revela-se incompleta para dar vazão a esse tipo de dinâmica normativa.

Por isso, os dados que influem na projeção de uma pessoa e na sua esfera relacional adéquam-se conceitualmente como um novo direito da personalidade⁸⁰. Alocar a proteção dos dados pessoais nessa categoria jurídica é uma construção dogmática necessária. Além de dar coerência normativa a uma série de faculdades jurídicas próprias desse direito (e.g., direito de acesso, correção, revisão de decisões automatizadas etc.), trata-se de um norte que facilita a sua interpretação e aplicação para não empolar a compreensão de seus conceitos basilares.

2.2.2 Conceito de dados pessoais: reducionista *versus* expansionista nas leis setoriais e na lei geral brasileira de proteção de dados pessoais⁸¹

O conceito de dados pessoais é um elemento central⁸² para que se aperfeiçoe a normatização sob análise, na medida em que se estabelecem os

New York: Information Science Reference, 2011. p. 34: "Moreover, this means that rules on the protection of personal data (defined as any information, truthful or not, relating to an identified or identifiable person) go clearly beyond the protection of privacy, covering also the protection (and promotion) of one's identity".

⁷⁸ Vide nota de rodapé 73, *supra*.

⁷⁹ CAPELO DE SOUSA, Rabindranath V. A. Op. cit. p. 245: "O bem da identidade reside, assim, na própria ligação de correspondência ou identidade do homem consigo mesmo e está, pois, ligado a profundas necessidades humanas, a ponto de o teor da convivência humana depender de sua salvaguarda em termos de reciprocidade".

⁸⁰ FINOCCHIARO, Giusella. *Privacy...* Op. cit., p. 5: "Il diritto alla protezione dei dati personali va collocato nell'ambito dei diritti della personalità. Si tratta, como é noto, di una categoria aperta di diritti caratterizzati dall'essere assoluti, indisponibili e imprescrittibili".

⁸¹ Uma parcela deste subcapítulo congrega parte das contribuições de outro trabalho de nossa autoria: BIONI, Bruno Ricardo. Xequê-Mate: o tripé de proteção de dados pessoais no xadrez das iniciativas legislativas no Brasil. Grupo de Pesquisa em Políticas Públicas para o Acesso à Informação/GPoPAI da Universidade de São Paulo. 2016. Disponível em: < https://www.academia.edu/28752561/Xequê-Mate_o_trip%C3%A9_de_prote%C3%A7%C3%A3o_de_dados_pessoais_no_xadrez_das_iniciativas_legislativas_no_Brasil >.

⁸² SCHWARTZ Paul; M. SOLOVE, Daniel J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *Review Law N.Y.U.*, p. 1816. Dis-

limites da própria tutela jurídica em questão. Em outras palavras, um dado que não avoque tal qualidade não poderia ser cogitado como um prolongamento da pessoa por lhe faltar tal centro de imputação.

Mutatis mutandis, seria a mesma lógica do fato jurídico⁸³. Não seria qualquer dado que teria repercussão jurídica, mas, somente, aquele que atraísse o qualificador pessoal.

De forma sistemática, o vocabulário para prescrever tal definição é composto por palavras que restringem ou alargam o *gargalo* dessa proteção. Há uma bipartição do seu léxico que ora retrai (reducionista), ora expande (expansionista), a moldura normativa de uma lei de proteção de dados pessoais.

Os quadros a seguir sintetizam qual é esse vocabulário e qual foi a opção adotada na LGPD e nas normas setoriais brasileiras de proteção de dados pessoais para, em seguida, proceder a uma análise ilustrativa das suas repercussões práticas.

Quadro 1 – Vocabulário analítico para a definição do conceito de dados pessoais

Expansionista	Reducionista
Pessoa identificável	Pessoa identificada
Pessoa indeterminada	Pessoa específica/determinada
Vínculo mediato, indireto, impreciso ou inexato	Vínculo imediato, direto, preciso ou exato
Alargamento da qualificação do dado como pessoal	Retração da qualificação do dado como pessoal

ponível em: <<http://scholarship.law.berkeley.edu/facpubs/1638>>: “PII is one of the most central concepts in privacy regulation. It defines the scope and boundaries of a large range of privacy statutes and regulations”.

⁸³ Fato jurídico é aquele que tem repercussão para o direito, produzindo efeitos jurídicos, criando, modificando e extinguindo relações jurídicas. Um fato sem tal adjetivação não tem repercussão e, portanto, não produz efeitos jurídicos, como é o exemplo da queda de uma folha seca da árvore que em nada tem relevância para a ciência jurídica (AMARAL, Francisco. *Direito...* Op. cit., p. 379-380).

Quadro 2 – Tabela analítica da definição de dados pessoais das propostas legislativas no Congresso Nacional

LGPD	Leis Setoriais de PDP	
	Decreto 8.771/2016	Lei 12.527/2011
Art. 5º, inciso I – dado pessoal: dado relacionado à pessoa natural identificada ou identificável	Art. 14, inciso I – dado pessoal: dado relacionado à pessoa natural identificada ou identificável , inclusive números identificativos, dados locacionais ou identificadores eletrônicos, quando estes estiverem relacionados a uma pessoa	Art. 4º, inciso IV – informação pessoal: aquela relacionada à pessoa natural identificada ou identificável
Expansionista (sem rol exemplificativo)	Expansionista (com rol exemplificativo)	Expansionista (sem rol exemplificativo)

A intelecção do conceito de dado pessoal e, por conseguinte, das estratégias regulatórias possíveis para a sua definição é algo fluido, que pode ser esclarecido a partir da dinâmica de conceitos básicos de sistemas de informação e de banco de dados. Somente, assim, o seu vocabulário ganhará uma análise mais concreta a demonstrar as diferenças e consequências práticas entre tais estratégias regulatórias distintas.

Para fins de ilustração, optou-se pelo exemplo de um banco de dados relacionais. Esse tipo de banco de dados é estruturado por tabelas⁸⁴, em que cada uma de suas colunas – que são chamadas de atributos⁸⁵ – é a maneira pela qual os dados são organizados. É a correlação entre as colunas e as linhas dessa tabela que empresta valor (significado) aos dados, permitindo que deles seja extraído algo inteligível (informação)⁸⁶.

Esse é um exercício didático para entender como os referenciais teóricos reducionista e expansionista distanciam-se completamente um do outro, apesar da proximidade léxica entres eles.

⁸⁴ MANNINO, Michael V. *Projeto...* Op. cit., p. 555.

⁸⁵ ROB, Peter. *Sistemas...* Op. cit., p. 4.

⁸⁶ Ibidem.

Tabela 1 – Base de dados relacionais

A) Nome	B) CPF	C) CEP	D) Idade	E) Classificação/ segmentação
1. Bruno Santos	123.456-77	04055-000	18	Jovem hipster
2. Bruno Santos	234.567-88	04055-111	17	Jovem poupador
3. Bruno Santos	345.678-99	04055-222	17	Jovem consumista
4. Bruno Souza	456.789-10	01201-000	65	Idoso com rentabilidade
5. Bruna Souza	567.891-01	04201-111	66	Idosa sem rentabilidade
6. Bruna Bioni	222.333-44	04201-222	70	Idosa com rentabilidade
7. Maria Silva	157.890-88	09201-000	40	Adulto desempregado
8. Maria Silva	666.666-66	09201-111	38	Adulto perfil executivo
9. Maria da Silva	987.354-22	09201-222	16	Jovem hipster

Hipótese 1 (reducionistas): a presença de homônimos não permitiria que houvesse a individualização precisa de uma parcela das pessoas inseridas no banco de dados acima, caso não houvesse outros dados – identificadores (únicos) como, por exemplo, o CPF. Somente mediante tal associação identificou-se exatamente o(s) “Bruno(s)” e a(s) “Maria(s)”, tornando-os pessoas identificadas, isto é, individualizadas de maneira precisa, exata e inequívoca.

Hipótese 2 (expansionistas): caso houvesse a eliminação da coluna “B”, haveria incerteza a respeito de qual dos “Brunos” seria consumista, poupador ou hipster (coluna “E”), já que não haveria outro dado – identificador (único) – para diferenciar cada um dos homônimos. E, o mesmo, com relação a qual Maria seria a desempregada ou a executiva. Tais informações não estariam relacionadas a pessoas identificadas, ainda que elas pudessem vir a ser a partir da sua localização geográfica (coluna “B”), por exemplo. Nesse caso, a incerteza gerada por um grupo de pessoas com o mesmo nome tende a ser eliminada pela agregação dessa outra informação, tornando as pessoas identificáveis. Há, em última análise, o potencial de individualizá-las.

Portanto, verificar se um dado pode ser adjetivado como pessoal é uma *análise contextual* que depende de qual tipo de informação pode ser extraída de uma base de dados. Essa análise circunstanciada pode ser mais dura ou mais flexível. Para os reducionistas, somente na hipótese “A” haveria dados pessoais. Para os expansionistas, as hipóteses “A” e “B” seriam abraçadas pelo conceito de dado pessoal.

2.2.3 Dados “anônimos” como a antítese de dados pessoais: o filtro da razoabilidade

A antítese do conceito de dado pessoal seria um dado anônimo, ou seja, aquele que é incapaz de revelar a identidade de uma pessoa. Diante do próprio significado do termo, anônimo seria aquele que não tem nome nem rosto⁸⁷.

⁸⁷ HOUAISS, Antônio; VILLAR, Mauro de Salles. Op. cit., p. 140-141.

Essa inaptidão pode ser fruto de um processo pelo qual é quebrado o vínculo entre o(s) dado(s) e seu(s) respectivo(s) titular(es), o que é chamado de anonimização⁸⁸. Esse processo pode se valer de diferentes técnicas que buscam eliminar tais elementos identificadores de uma base de dados⁸⁹, variando entre: a) supressão; b) generalização; c) randomização e; d) pseudoanonimização⁹⁰.

Foge ao escopo deste trabalho abordar cada uma das citadas técnicas do processo de anonimização. Vale a pena, no entanto, abordar as duas primeiras para ilustrar a sua dinâmica e pavimentar a via de acesso para se debruçar sobre as implicações normativas de uma eventual dicotomia entre dados anônimos (anonimizados) e dados pessoais.

Retomando o exemplo de base de dados relacionais estruturada, deve-se identificar quais elementos poderiam ser modificados – suprimidos ou generalizados – para que o seu grau de identificabilidade seja eliminado ou reduzido:

- a) **supressão do CPF:** por ser um identificador capaz de diferenciar até mesmo pessoas homônimas, sendo um identificador único; logo, a sua disponibilização, ainda que parcial – e.g., cinco primeiros dígitos –, não seria prudente;
- b) **generalização do nome completo:** constaria apenas o prenome, desde que fosse observado que os nomes da base de dados não são comuns. O objetivo é evitar que um nome possa ser atribuído a um indivíduo em específico;
- c) **generalização da localização geográfica:** em vez de disponibilizar o número completo do CEP, seriam divulgados apenas os seus primeiros dígitos. Assim, haveria uma localização menos detalhada, a fim de quebrar o vínculo de identificação desta informação com um sujeito;

⁸⁸ DONEDA, Danilo. *Da privacidade...* Op. cit., p. 44-43.

⁸⁹ COUNCIL OF EUROPE. *Handbook on European Data Protection Law*. Luxembourg: Publications Office of the Europe Union, 2014. Disponível em: <http://www.echr.coe.int/Documents/Handbook_data_protection_ENG.pdf>.

⁹⁰ Para muitos, a pseudoanonimização não é considerada uma técnica de anonimização. Isso porque se substituem, apenas, os identificadores diretos – e.g., nome, CPF etc. – por pseudônimos – e.g., números aleatórios, de modo que a pessoa permanece sendo identificável em razão de tais pseudônimos serem um retrato detalhado indireto delas (WP 29, 2014, p. 20).

d) **generalização da idade:** em vez de divulgar a idade exata, seria divulgada a faixa etária para viabilizar a categorização dos indivíduos como jovens, adultos ou idosos (coluna "E") e, por outro lado, inviabilizar a sua individualização, dado o universo de pessoas que se enquadram naquela mesma faixa etária.

Tabela 2 – Base de dados relacionais anonimizada

A) Nome	B) CPF	C) CEP	D) Faixa etária	E) Classificação/segmentação
1. Bruno dos Santos	123.456-77	04055-000	18>	Jovem hipster
2. Bruno dos Santos	234.567-88	04055-111	18>	Jovem poupador
3. Bruno dos Santos	345.678-99	04055-222	18>	Jovem consumista
4. Bruna Souza	456.789-10	01201-000	60<	Idoso com rentabilidade
5. Bruna Souza	567.891-01	04201-111	60<	Idosa sem rentabilidade
6. Bruna Schonber	222.333.44-55	04201-222	60<	Idosa com rentabilidade
7. Maria Silva	157.890.88-66	09201-000	18<	Adulto desempregado
8. Maria Silva	666.666.66-66	09201-111	18<	Adulto perfil executivo
9. Maria Sóstenes	987.354.22-99	09201-222	18>	Jovem hipster

Com maior ou menor grau de intensidade – e.g., supressão ou generalização – nota-se um método cujo mote é gerenciar circunstancialmente a *identificabilidade* de uma base de dados. As características de cada dado e a percepção de eles estarem inseridos em uma gama de informações devem orientar tal análise.

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

Amarrar o conceito teórico de dados anônimos a uma análise contextual, com os olhos voltados para a irreversibilidade do processo de anonimização, joga luz diretamente sobre o fator problemático dessa proposição: o seu caráter elusivo ou mesmo a sua impossibilidade teórica⁹¹.

⁹¹ TEIXEIRA, Lucas. Teoricamente impossível: problemas com a anonimização de dados pessoais. Disponível em: <<https://antivigilancia.org/pt/2015/05/anonimizacao-dados-pessoais/>>.

Torna-se cada vez mais recorrente a publicação de estudos que demonstram ser o processo de anonimização algo falível. A representação simbólica de que os vínculos de identificação de uma base de dados poderiam ser completamente eliminados, garantindo-se, com 100% (cem por cento) de eficiência, o anonimato das pessoas, é um mito⁹².

Os pesquisadores Arvind Narayanan e Vitaly Shmatikov têm se destacado nessa área. Eles reidentificaram diversas bases de dados anonimizados, como, por exemplo, a de um famoso provedor de aplicação de Internet – o caso Netflix Prize.

Nesse caso, os referidos pesquisadores desenvolveram um algoritmo que calculava dentre outras coisas: a) quantos *bits* de informação seriam necessários para reverter o processo de anonimização; b) qual seria o melhor critério para a escolha de uma informação auxiliar – i.e., uma outra base de dados – a ser agregada para reverter o processo de anonimização; c) uma métrica sobre a probabilidade de acerto da reidentificação, evitando-se “falsos positivos” – i.e., a vinculação errônea de indivíduos aos dados desanonimizados⁹³.

À época, a maior provedora de *streaming* de filmes do mundo criou um concurso, cujo desafio era melhorar o seu algoritmo de sugestão de filmes. Então, a Netflix disponibilizou a sua base de dados com todas as avaliações dos filmes de seu catálogo do período de 1998 a 2005, suprimindo os nomes dos usuários avaliadores e deixando somente a data e a nota da avaliação.

A fim de tornar tal processo de reidentificação mais robusto, a Netflix se utilizou da técnica de randomização. Ela alterou algumas datas e *ratings* das avaliações dos seus consumidores, o que aumentaria o risco de “falsos positivos”, a não ser pelo fato de o algoritmo dos pesquisadores ter sido projetado para isso.

Os pesquisadores “rodaram” tal algoritmo na base de dados disponibilizada, descobrindo que seria necessário entre 3 (três) e 19 (dezenove) *bits* de informação para reverter o processo de anonimização. Esse *pool* de informações necessário estava publicamente disponível e acessível na Internet Movies Databases/IMDB⁹⁴.

O IMDB é um *website* onde as pessoas compartilham suas impressões sobre filmes, utilizando-se, na maioria das vezes, dos seus nomes reais. Desta forma, os pesquisadores “cruzaram” essas informações com a base de dados da

⁹² NARAYANAN, Arvind; SHMATIKOV, Vitaly. Myths and Fallacies of “Personally Identifiable Information”. *Communications of the ACM*, v. 53, n. 06, p. 24, June 2010. Disponível em: <www.cs.utexas.edu/~shmat/shmat_cacm10.pdf>.

⁹³ NARAYANAN, Arvind; SHMATIKOV, Vitaly. Robust De-anonymization of Large Sparse Datasets, p. 6. Disponível em: <https://www.cs.utexas.edu/~shmat/shmat_oak-08netflix.pdf>.

⁹⁴ NARAYANAN, Arvind; SHMATIKOV, Vitaly. Robust... Op. cit., p. 12.

Netflix, correlacionando as datas das avaliações dos filmes e seus respectivos scorings. Assim, a peça faltante do quebra-cabeça – a identidade dos usuários da Netflix – foi desvendada com base nos nomes contidos nas avaliações do IMDB⁹⁵.

O exemplo em questão é simbólico, pois sublinha o “calcanhar de Aquiles” dos dados anônimos. Sempre existirá a possibilidade de uma base de dados anonimizada ser agregada a outra para a sua reidentificação⁹⁶⁻⁹⁷.

Por isso, via de regra, prevalece conceito *expansionista*⁹⁸ pelo qual dado pessoal equivale a uma informação que, direta ou indiretamente, identifica um sujeito. Essa definição abraça, portanto, mesmo as informações que têm o potencial de identificar alguém, ainda que de maneira remota⁹⁹.

Desde a década de 1980, organismos internacionais, blocos econômicos, países norte-americanos e latino-americanos conceituam dados pessoais como uma informação a respeito de uma pessoa identificada ou identificável¹⁰⁰.

OCDE	CoE	Canadá	Argentina	APEC	UE
“personal data” means any information relating to an identified or <u>identifiable</u> individual (data subject)	“personal data” means any information relating to an identified or <u>identifiable</u> individual (“data subject”)	personal information means information about an <u>identifiable</u> individual.	Datos personales: Información de cualquier tipo referida a personas físicas o de existencia ideal determinadas o <u>determinables</u> .	Personal information means any information about an identified or <u>identifiable</u> individual.	“personal data” means any information relating to an identified or <u>identifiable</u> individual (“data subject”);
<u>Expansionista</u>	<u>Expansionista</u>	<u>Expansionista</u>	<u>Expansionista</u>	<u>Expansionista</u>	<u>Expansionista</u>
1980	1981	2000	2000	2005	1995/2016

⁹⁵ Ibidem, p. 15.

⁹⁶ OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization (August 13, 2009). *UCLA Law Review*, v. 57, p. 1749, 2010; *Colorado Law Legal Studies Research Paper* n. 9-12. Disponível em: <<http://ssrn.com/abstract=1450006>>.

⁹⁷ NARAYANAN, Arvind; SHMATIKOV, Vitaly. Robust... Op. cit., p. 4.

⁹⁸ SCHWARTZ Paul M.; SOLOVE, Daniel J. The PII... Op. cit., p. 1874.

⁹⁹ Vejam-se, no entanto, nossas ressalvas feitas em outro trabalho em que se leva em consideração o critério de razoabilidade para transformar um dado anônimo em dado pessoal (processo de reidentificação). Nele, nós concluímos ser possível a convivência entre um conceito *expansionista* e o de dado anonimizado: BIONI, Bruno. Xequemate... Op. cit., p. 30-36.

¹⁰⁰ Art. 5º, I: “dado relacionado à pessoa natural identificada ou identificável, inclusive números identificativos, dados locacionais ou identificadores eletrônicos quando estes estiverem relacionados a uma pessoa”.

A proteção dos dados pessoais, como um novo direito da personalidade, dirige-se a todo e qualquer dado em que se denote o *prolongamento* de um sujeito. Dados pessoais não se limitam, portanto, a um tipo de projeção *imediata*, mas, também, a um referencial *mediato* que pode ter *ingerência* na esfera de uma pessoa¹⁰¹.

Por essa lógica, qualquer dado pessoal anonimizado detém o *risco inerente* de se transmutar em um dado pessoal¹⁰². A agregação de diversos “pedaços” de informação (dados) pode revelar (identificar) a imagem (sujeito) do quebra-cabeça, a qual era até então desfigurada (anônimo) – o chamado efeito mosaico.

Por isso, em princípio, eventual dicotomia entre dados pessoais e dados anônimos só guardaria coerência junto ao conceito reducionista de dados pessoais. Isso porque dados anônimos não são dados relacionados a uma pessoa identificada, demandando a reversão do processo de anonimização para se chegar aos respectivos titulares, sendo a sua identificabilidade remota (identificável) e não imediata (identificada).

Dessa forma, leis que adotam o conceito expansionista de dados pessoais e, ao mesmo tempo, estabelecem uma dicotomia deste com dados anônimos correriam o risco de ser tautológicas. Isso porque haveria uma *redundância normativa*, já que dados anônimos seriam, em última análise, potencial e provavelmente, dados relacionados a uma pessoa identificável.

Para não gerar tal incoerência, a única saída foi a adoção de um “filtro” que delimitasse a *elasticidade* do conceito expansionista – neste caso o termo identificável –, sob pena de a fronteira entre dados pessoais e dados anônimos ser sempre transponível.

E, nesse sentido, o direito comunitário europeu¹⁰³ e a LGPD¹⁰⁴ valeram-se do critério da razoabilidade para delimitar o espectro do conceito

¹⁰¹ FINOCCHIARO, Giusella. *Privacy...* Op. cit., p. 55: “Il concetto chiave, tanto nella definizione di dato personale, quanto in quella di dato anonimo, è quello di riferibilità. Se le informazioni possono essere riferite ad un soggetto, i dati non sono anonimi, ma personali e, conseguentemente, deve applicarsi la normativa in materia di protezione dei dati personali”.

¹⁰² TENE, Omer. Privacy law’s midlife crisis: a critical assessment of the second wave of global privacy laws. *Ohio State Journal*, v. 74, 2013, p. 1.242.

¹⁰³ A Diretiva 95/46 e a sua proposta de regulamentação adotam os conceitos de razoabilidade, respectivamente, nas considerandas 26 (vinte e seis) e 23 (vinte e três).

¹⁰⁴ Na definição de dados anônimos, de anonimização, bem como no dispositivo que prevê em quais hipóteses dado anonimizado pode ser considerado como dado pessoal, a LGPD faz alusão ao termo razoável(is) – respectivamente, arts. 5º, II e III, e 18.

expansionista de dados pessoais. Não basta a mera possibilidade de que um dado seja atrelado a uma pessoa para atrair o termo identificável¹⁰⁵. Essa vinculação deve ser objeto de um “esforço razoável”¹⁰⁶, sendo esse o perímetro de elasticidade do conceito de dado pessoal como aquele relacionado a uma pessoa identificável.

A *contrario sensu*, se para a correlação entre um dado e uma pessoa demanda-se um esforço fora do razoável, não há que se falar em dados pessoais. Nessa situação, o dado é considerado como anônimo, uma vez que o “filtro da razoabilidade” barra o seu enquadramento como aquele relacionado a uma pessoa identificável¹⁰⁷.

A própria LGPD estabelece critérios para interpretar esse conceito equívoco de “razoabilidade” em seu art. 18, § 1º: “A determinação do que seja razoável deve levar em consideração fatores objetivos, tais como custo e tempo necessários para reverter o processo de anonimização, de acordo com as tecnologias disponíveis, e a utilização exclusiva de meios próprios”.

Trata-se de uma estratégia normativa alinhada à premissa de que os dados anônimos seriam sempre passíveis de reversão. Daí por que o critério da razoabilidade nada mais é do que uma diretriz acerca do que venha a ser um *risco aceitável-tolerável*¹⁰⁸ em torno da reversibilidade do processo de anonimização, a fim de que os dados anonimizados estejam fora do conceito de dados pessoais.

Com isso, há coerência em se estabelecer conceitos diferentes para tais espécies de dados, sobretudo sob o ponto de vista de uma dicotomia mutuamente excludente entre eles, que é delimitada pelo fator da

¹⁰⁵ ARTICLE 29, Data Protection Working Party. Opinion 04/2007 on the concept of personal data, p. 1749. Disponível em: <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>. p. 15.

¹⁰⁶ Essa é exatamente a terminologia utilizada pelo art. 18, *caput*, da LGPD: “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 revertidos”.

¹⁰⁷ Ibidem, p. 21.

¹⁰⁸ Sobre a estratégia regulatória baseada no risco e, em particular, relacionada ao conceito de dado pessoal e dado anonimizado, veja-se: RUBINSTEIN, Ira; HARTZOG, Woodrow. Anonymization and Risk (August 17, 2015). 91 *Washington Law Review*, 703, 2016; NYU School of Law, *Public Law Research Paper* n. 15-36. Disponível em: <<https://ssrn.com/abstract=2646185>>.

razoabilidade¹⁰⁹. Do contrário, repita-se, haveria uma redundância normativa, na medida em que dados anônimos – sem o critério da razoabilidade – seriam sempre enquadrados dentro do conceito de dado pessoal, como aquele relacionado a uma pessoa identificável.

2.2.4 A importância pragmática da alocação dogmática da proteção de dados pessoais como um novo direito da personalidade: perfis comportamentais e decisões automatizadas (arts. 12, § 2º, e 20 da LGPD)

Certa vez, um engenheiro do Google teria dito que eles não coletam informações associadas aos nomes das pessoas, pois isso geraria desinformação – “ruído” nas palavras dele¹¹⁰. Em outra oportunidade, o então chefe de assuntos de privacidade do Facebook, Erin Egan, afirmou paradoxalmente que, apesar de a rede social fornecer publicidade com base na identidade dos seus usuários, isso não significaria que eles sejam pessoas identificáveis¹¹¹. Em uma série de reportagens sobre a técnica de “Privacy Differential” da Apple, a revista *Wired* é provocativa ao dizer, paradoxalmente, que essa técnica “coleta dados sobre você”, mas não muito “bem sobre você”¹¹².

De fato, o modo pelo qual a Internet funciona não torna necessário saber a identidade do usuário para lhe direcionar um conteúdo, ou, mesmo, sujeitá-lo a um processo de decisão automatizada. Basta lhe atribuir um identificador eletrônico único que permita separá-lo dos milhões de usuários da rede, como, por exemplo, em relação ao computador ao qual ele está conectado, o que é feito por meio do número de conexão a ele atribuído (o chamado protocolo de endereço – IP; vide subcapítulo 2.3.2 *infra*)¹¹³.

¹⁰⁹ Sobre as disputas interpretativas em torno do conceito jurídico indeterminado de razoabilidade, veja-se: BIONI, Bruno Ricardo. Xequemate... Op. cit., p. 34-35.

¹¹⁰ HARDY, Quentin. *Rethinking Privacy in an Era of Big Data*. Disponível em: <http://bits.blogs.nytimes.com/2012/06/04/rethinking-privacy-in-an-era-of-big-data/?_r=0>.

¹¹¹ DEVRIES, Jennifer Valentino; VINE, Jeremy Singer. *They Know What You're Shopping For*. Disponível em: <<http://www.wsj.com/articles/SB10001424127887324784404578143144132736214>>.

¹¹² Disponível em: <<https://www.wired.com/2016/06/apples-differential-privacy-collecting-data/>>.

¹¹³ O Marco Civil acabou por definir o conceito de protocolo IP em seu art. 5º, III: “III – endereço de protocolo de internet (endereço IP): o código atribuído a um terminal de uma rede para permitir sua identificação, definido segundo parâmetros internacionais”.